STRUCTURAL ANALYSIS OF LIPOPROTEIN MODIFICATIONS IN MYCOBACTERIA AND OTHER GC-RICH GRAM-POSITIVE BACTERIA

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Juliane K. Brülle

aus
Deutschland

Promotionskomitee

Prof. Dr. Erik C. Böttger (Vorsitz)
Prof. Dr. Peter Sander (Leitung der Dissertation)
Prof. Dr. Leo Eberl

Zürich 2015

To my mum and dad

“A man who is not confused is not well informed”

(Irish saying)

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to the many people who contributed to the final success of this thesis.

First I would like to thank my supervisors Prof. Dr. Erik C. Böttger and Prof. Dr. Peter Sander for giving me the opportunity to work in their institute. Peter, I am grateful for your support and inspiration for my work and your constructive critisism in our discussions. I also want to thank Prof. Dr. Leo Eberl from the Department of Microbiology (UZH) for joining my thesis committee.

Special thanks go to Yolanda Joho-Auchli from the Functional Genomics Center Zurich for her helpfulness and the performance of the MALDI-TOF/TOF analyses.

I further want to thank all my former and present colleagues of the institute, especially the members of the FoSa group: Petra, Andreas, Tobi, Peter K., Silke, Michael M., Corrado, Janine, Reto, Agnese, Beat, Bernadette, Katja, Anna K., Michael D. M. and Anna R. You all contributed to a comfortable and friendly atmosphere in the lab and I am grateful for the social events we had together.
Sincere thanks are given to Thomas Grau, for supervising my diploma thesis and for his great advice and help. Without him my first publication would not have been possible.

I would further like to thank my closest friends Anna, Dori, Alexa, Caro, and Mads for their long-lasting support and their patience. Your distractions and exhilaration always provided me with new power and motivation. Very warm thanks go to Petra and Janis, who have always been there for me and who lent me a sympathetic ear at any time. I also want to thank Ulli & Michael and Jürgen & Gislinde for accompanying my way and for their pleasant help.

My deepest gratitude goes to my family, especially my mum and dad who unswervingly supported me throughout my studies and who always believed in me and what I was doing.

Last but not least, I want to thank Matthias Thöny for his trust in my abilities and his encouraging words, keeping me from becoming too desperate in bad times.

Thank you all!

TABLE OF CONTENTS

SUMMARY

Tuberculosis is one of the most pernicious diseases in the world. With more than 1.3 million death per year, tuberculosis ranks as the second leading cause of death due to a single infectious agent. *Mycobacterium tuberculosis*, the causative agent of tuberculosis is a slow-growing, G+C-rich Gram-positive bacterium. Mycobacteria in general are characterized by a unique and complex cell envelope which forms a permeability barrier conferring natural resistance to many chemical compounds and therapeutic agents. Therefore, tuberculosis treatment is difficult and lengthy, but especially the increased occurance of antibiotic resistant bacteria or bacteria resistant to almost every available drug, exacerbates the fight against the disease. To substantially improve this situation, new anti-tuberculosis drugs or potential drug targets, new vaccines and a detailed understanding of host-pathogen interactions are needed.

Lipoproteins are secreted proteins N-terminally modified with fatty acids by the consecutive action of preprolipoprotein diacylglyceryl transferase (Lgt), prolipoprotein signal peptidase (LspA) and apolipoprotein *N*-acyltransferase (Lnt). Mycobacterial lipoproteins are immunodominant antigens, feature many diverse functions and are involved in several biological processes correlated with virulence. Thus, lipoproteins are candidates for novel drug targets or subunit vaccines.

The aim of this study was the molecular characterization of the membrane anchor of lipoproteins 1) in the fast-growing *Mycobacterium smegmatis*, 2) the slow-growing *Mycobacterium bovis* BCG and 3) in the related *Streptomyces scabies.* The identification of enzymes involved in the synthesis of the membrane anchor was an additional aim.

M. tuberculosis lipoproteins were heterologously expressed in *M. smegmatis* and *M. bovis* BCG, purified and analysed by MALDI-TOF/TOF analyses. The structure of the lipid modification was solved at the molecular level and determined as thioether-linked diacylglyceryl residue and an amide-linked third fatty acid. Palmitic, oleic and tuberculostearic acid were found as ester-linked moieties in the diacylglyceryl residue, while palmitic acid was found as fatty acid for the *N*-acyl. Aside from that, tuberculostearic acid was identified as substrate for *N*-acylation in *M. bovis* BCG. Isogenic knockout mutants in the gene homologous to *Escherichia coli lnt* have been generated to investigate *N*-acylation. ORF BCG_2070c was identified as apolipoprotein *N*-acyltransferase in *M. bovis* BCG.

Analysis of lipoproteins expressed in *S. scabies* showed a similar acylation pattern as in mycobacteria and demonstrated Lnt function in a monoderm GC-rich Gram-positive bacterium. Therefore, lipoprotein synthesis by Lgt, LspA and Lnt modification seems to be a conserved mechanism in bacteria.

ZUSAMMENFASSUNG

Tuberkulose gilt als eine der verheerendsten Krankheiten der Welt. Mit mehr als 1.3 Millionen Todesfällen pro Jahr ist Tuberkulose die zweithäufigste Todesursache unter den Infektionskrankheiten. Der Erreger der Tuberkulose, *Mycobacterium tuberculosis*, ist ein langsam wachsendes, Gram-positives Bakterium, dessen Genom einen hohen G+C Anteil aufweist. Mykobakterien im Allgemeinen sind durch eine einzigartige und komplexe Zellhülle gekennzeichnet, welche eine Permeabilitätsbarriere darstellt und somit natürliche Resistenz gegen chemische Substanzen und therapeutische Arzneimittel verleiht. Dies macht die Tuberkulosebehandlung schwierig und langwierig, aber besonders die wachsende Anzahl antibiotikaresistenter Bakterien und solcher, die gegen annähernd alle vorhandenen Medikamente resistent sind, erschwert den Kampf gegen die Krankheit. Um diese Situation erheblich zu verbessern werden dringend neue anti-tuberkulöse Medikamente, neue Impfstoffe und ein detailliertes Wissen über die Wirt-Pathogen Interaktionen benötigt.

Lipoproteine sind sekretierte Proteine, welche durch die aufeinanderfolgende Aktion von Pre-prolipoprotein Diacylglyceryltransferase (Lgt), Prolipoprotein Signalpeptidase (LspA) und Apolipoprotein *N*-acyltransferase (Lnt) N-terminal mit Fettsäuren modifiziert werden. Mykobakterielle Lipoproteine sind immunodominante Antigene. Sie weisen viele verschiedene Funktionen auf und sind in diverse biologische Prozesse involviert, welche mit der Virulenz in Verbindung gebracht werden. Aus diesem Grund stellen Lipoproteine Kandidaten als Zielstrukturen für neue Medikamente oder Impfstoffe dar.

Das Ziel dieser Arbeit war die molekulare Charakterisierung des Membran-Ankers von Lipoproteinen 1) im schnell-wachsenden *Mycobacterium smegmatis*, 2) im langsam-wachsenden *Mycobacterium bovis* BCG und 3) im verwandten *Streptomyces scabies.* Ein weiteres Ziel war die Indentifizierung der Enzyme, welche an der Synthese des Membran-Ankers beteiligt sind.

M. tuberculosis Lipoproteine wurden heterolog in *M. smegmatis* und *M. bovis* BCG exprimiert, aufgereinigt und per MALDI-TOF/TOF Analyse analysiert. Die Struktur des Lipidankers wurde bis zur molekularen Ebene aufgeklärt und als thioether-gebundener Diacylglyceryl Rest und amid-gebundene dritte Fettsäure entschlüsselt. Palmitat, Oleat und Tuberculostearat wurden als ester-gebundene Strukturen des Diacylglyceryl Rests und Palmitat als Fettsäure für das *N*-acyl gefunden. Desweiteren wurde Tuberculostearat als Substrat für die *N*-acylierung in *M. bovis* BCG identifiziert. Zur Untersuchung der *N*-acylierung wurden isogene Deletionsmutanten im homologen Gen zum *Escherichia coli lnt* generiert. Dabei wurde der offene Leserahmen BCG_2070c als Apolipoprotein-*N*-acyltranferase in *M. bovis* BCG identifiziert.

Die Analyse von Lipoproteinen aus *S. scabies* ergab ein ähnliches Acylierungsmuster wie in den Mykobakterien und bestätigte somit eine Lnt Funktion auch in monodermen G+C-reichen Gram-positiven Bakterien. Die Lipoprotein-Modifizierung mit drei Fettsäuren durch Lgt, LspA und Lnt, kann daher als konservierter Mechanismus in Bakterien angesehen werden.

INTRODUCTION

„…erst durch die Entdeckung des Tuberkelbazillus wurde die Ätiologie der Tuberkulose auf eine sichere Grundlage gestellt und die Überzeugung gewonnen, dass dieselbe eine parasitäre, d. h. eine ansteckende, aber auch vermeidbare Krankheit ist. Schon bei den ersten Mitteilungen über die Ätiologie der Tuberkulose habe ich auf die Gefahren hingewiesen, welche durch die Verbreitung der bazillenhaltigen Absonderungen der Schwindsüchtigen entstehen, und habe dazu aufgefordert, prophylaktische Massregeln gegen die Seuche zu ergreifen. Aber meine Worte sind unbeachtet geblieben. Es war eben noch zu früh, und sie konnten deswegen noch keinem vollen Verständnis begegnen. Es ging damit wie bei so vielen ähnlichen Gelegenheiten in der Medizin, wo es auch langer Zeit bedurft hat, ehe alte Vorurteile überwunden und die neuen Tatsachen von den Ärzten als richtig anerkannt wurden. Aber ganz allmählich hat sich dann die Erkenntnis von der ansteckenden Natur der Tuberkulose verbreitet und immer tiefere Wurzeln gefasst, und je mehr die Überzeugung von der Gefährlichkeit der Tuberkulose sich Bahn brach, umso mehr hat sich auch die Notwendigkeit aufgedrängt, sich dagegen zu schützen.“

Dr. Robert Koch „Über den derzeitigen Stand der Tuberkulosebekämpfung“ Nobel-Vorlesung, gehalten in Stockholm am 12. Dezember 1905. Deutsche Medizinische Wochenschrift, 1906, Nr. 3

Tuberculosis is an ancient scourge. It has plagued mankind worldwide throughout much of known human history. Among the early record about tuberculosis, there are 3300 years old written texts from India describing the disease. In Egypt, tuberculosis can be dated back to even more than 5000 years ago, where typical skeletal abnormalities of tuberculosis have been found in mummies. Pulmonary tuberculosis was well known in ancient Greece. Hippocrates (460-370 BC) was the first describing it as “phthisis” or consumption and as the most considerable of the diseases. During the 18th and 19th centuries tuberculosis reached epidemic proportions in Europe and North America so that it was referred to as the “White Plague”. Among the thousands of victims it claimed were prominent people such as Friedrich Schiller, Frédéric Chopin, Charlotte Brontë, Lord Byron or Franz Kafka. Then the history of tuberculosis was changed. The identification of *Mycobacterium tuberculosis* as the etiologic agent by Robert Koch in 1882, the initiation of *Mycobacterium bovis* BCG by Albert Calmette and Camille Guérin as the first tuberculosis vaccine applied to a human in 1921 and the discovery of antibiotics such as streptomycin in 1944 and isoniazid in 1952 reduced the

death toll from tuberculosis for decades. Nevertheless, tuberculosis remains as deadly as it was (Daniel, 2006; Gengenbacher & Kaufmann, 2012).

Tuberculosis

Tuberculosis is still one of the most devastating diseases in the world. With 1.3 million deaths occuring in 2012 it is the second leading cause of death from a single infectious agent, after the human immunodeficiency virus (HIV) (http://www.who.int/tb/publications/factsheet_global.pdf?ua=1). The causative agent of tuberculosis, *M. tuberculosis* is primarily transmitted via the respiratory route and establishes its infection usually in the lungs, although most organs can be infected. Most people who come in contact with the pathogen remain clinically asymptomatic and develop a latent infection. It is estimated that one third of the world's population is latently affected (about 2 billion people). Only 5-10% of the infected individuals ever develop an active tuberculosis, however, with the number of 8.6 million new active cases in 2012 it is higher than at any other time in history (http://www.who.int/tb/publications/factsheet_global.pdf?ua=1).
The risk of resuscitation and transition from latent to active disease is increased by co-infection with HIV about 30 times compared to that in HIV-negative people (http://www.who.int/tb/challenges/hiv/tbhiv_factsheet_2013_web.pdf?ua=1). Other immuno-compromising treatments like anti-tumor necrosis factor therapy for chronic inflammatory diseases or factors like diabetes, alcoholism or poor nutrition are additional factors favouring for reactivation (Gengenbacher & Kaufmann, 2012).

Principally, the enormous success of *M. tuberculosis* is attributed to several capacities: the bacterium evades its elimination by innate and acquired host immune response by persisting inside macrophages, it is equipped with a thick and waxy cell wall which protects the pathogen against physical or chemical stress, it is naturally resistant to several antibiotics and slow-growing which makes antibiotic treatment difficult and longsome (Ducati *et al.*, 2006). In addition, the HIV/AIDS pandemic over the past few decades and the occurrence of drug-resistant strains represent a devastating problem in treatment of tuberculosis and account for its worldwide spread.

Co-infection of tuberculosis with HIV is the major risk for the conversion of latent into active disease. With one-third, tuberculosis is the most common presenting illness among the 35.3

million people living with HIV worldwide. In 2012, 320.000 people died from HIV-associated tuberculosis, which makes it the leading cause of death among HIV-infected individuals (http://www.who.int/tb/challenges/hiv/tbhiv_factsheet_2013_web.pdf?ua=1). The two infections mainly intersect in the world's poorest countries and where AIDS is prevalent, namely in sub-Saharan Africa and Asia, thereby highly magnifying the death-toll (Dietrich & Doherty, 2009).

M. tuberculosis is naturally resistant to several antibiotics. The resistant phenotypes can be attributed to the unusual structure and low permeability of the mycobacterial cell wall, the presence of drug efflux systems or drug-modifying enzymes (Niederweis *et al.*, 2010). In contrast, acquired resistance to anti-tuberculosis drugs primarily arises due to inadequate drug prescription, poor drug quality and non-compliance of the patients (Abubakar *et al.*, 2013). Over the past two decades acquired drug-resistance mechanisms led to the worldwide emergence of multidrug-, then extensively-, and, most recently, totally drug-resistant strains of *M. tuberculosis* (Migliori et al., 2007; Velayati et al., 2009). Tuberculosis caused by these resistant strains is much more difficult, cost-intensive and even longer to treat and threatens to undermine tuberculosis control (Abubakar *et al.*, 2013).

The standard treatment for drug-susceptible tuberculosis consists of a combination of the four first line drugs (isoniazid, rifampicin, pyrazinamide and ethambutol) for an initial phase of two months. In this phase, all replicating bacteria are eradicated, the sputum becomes sterile and the patient is no longer infectious. In the following four month phase a regimen of two first-line drugs, usually isoniazid and rifampicin, is given to eradicate the remaining non-replicating bacteria. The therapy of multidrug-resistant tuberculosis (MDR-TB), which is defined as tuberculosis caused by bacteria that are resistant to at least the two most effective first-line drugs, isoniazid and rifampicin (Böttger & Springer, 2008) is considerably longer (18-24 months) and based on the use of less effective second-line drugs that are associated with toxic effects (Lawn & Zumla, 2011). In 2012, there were more than 450.000 new multidrug-resistant cases in the world, with more than one half of these cases occurring in China, India and the Russian Federation. In recent years, *M. tuberculosis* strains additionally became resistant to an increasing number of drugs. Extensively drug-resistant (XDR) *M. tuberculosis* shows resistance to isoniazid and rifampicin as well as resistance to any fluoroquinolone and any of the injectable second-line drugs (amikacin, kanamycin and capreomycin) (http://www.who.int/tb/challenges/mdr/mdr_tb_factsheet.pdf?ua=1). Totally

drug-resistant tuberculosis bacteria are resistant to almost all antituberculosis drugs (Migliori *et al.*, 2007; Velayati *et al.*, 2009). Due to the resistance of XDR-*M. tuberculosis* to the most effective first- and second-line medications, patients with XDR-tuberculosis are often left with less-effective treatment options and it is not uncommon that people die with XDR-TB even after starting treatment (Abubakar *et al.*, 2013).

In order to contain the pandemic spread of tuberculosis, there is a strong need for an universally effective vaccine. The only licensed current vaccine, Bacille Calmette-Guérin (BCG) is an attenuated strain of *M. bovis*, the agent that causes tuberculosis in cattle. Albert Calmette and Camille Guérin developed BCG by 230 continuous *in vitro* passages in which the originally virulent *M. bovis* had lost its pathogenicity by undefined genetic alterations. The first dose was given to an infant in 1921, up to date, the vaccine has been administered to at least four billion individuals, making it the most widely used vaccine in the world (Weiner & Kaufmann, 2014). However, it has done little to satisfy the control of tuberculosis. Despite clear beneficial effects against tuberculosis in childhood, protection induced by BCG can wane within a number of years and thus, the protection from adult pulmonary tuberculosis is insufficient and varies from 80% to 0% (Gupta *et al.*, 2007). In addition, BCG is a pre-exposure vaccine and fails to provide protective immunity in people pre-sensitized to mycobacteria by *M. tuberculosis* infection or prior BCG vaccination. This means that BCG cannot be used as a booster vaccine to overcome the transient effect of the vaccination administered soon after birth (Dietrich & Doherty, 2009). Hence, new tuberculosis vaccines are urgently needed. There have been recent advantages in development of new vaccines with some candidates already undergoing clinical trials. The majority of these vaccines aim at prevention of active tuberculosis. These include pre-exposure prime vaccines, which could replace BCG, or pre-exposure booster vaccines given during childhood in addition to BCG prime vaccination. Post-exposure booster vaccines are designed for adolescents or adults with latent tuberculosis infection who had been vaccinated with BCG in childhood. Some therapeutic vaccines are tested for their potential as adjuvants in chemotherapy during active tuberculosis in adults (Weiner & Kaufmann, 2014). The ideal vaccine would be counteracting *M. tuberculosis* in all different stages of the infection, thereby being effective against active disease and latent infection. However, no such “multistage” vaccine is expected to become available until 2020 (Dietrich & Doherty, 2009).

Pathogenesis

The infection with the pathogen *M. tuberculosis*, the causative agent of tuberculosis follows a well-defined sequence of events. The infectious bacilli are aerogenically transmitted and after inhalation they are phagocytosed by alveolar macrophages in the lung and induce a localized inflammatory response (Ducati *et al.*, 2006). Only 1-10 bacilli are necessary to cause an infection (Bloom & Murray, 1992). Once the bacterium parasitizes the macrophage, the host usually is not able to eliminate, but to contain the pathogen by remodelling the site of infection into granuloma. These lesions, also called tubercles, which define the disease, contain infected macrophages in the center. This restricts growths and prevents the dissemination of the bacteria marking the end of the primary infection. Nevertheless, there are also some individuals that progress rapidly into active disease after initial infection, and, also some people that are resistant to *M. tuberculosis* or may have cleared infection through an effective innate immune response (O'Garra *et al.*, 2013).

During the following latent phase of the disease, there are no overt signs of tuberculosis like the typical cough and there is no infectious risk for others. The bacteria coexist within the host in a dormant form of infection, though still representing a large bacterial reservoir among infected individuals. The survival inside the macrophages is an attribute to the ability of *M. tuberculosis* to manipulate the phagosome in which it resides. The bacillus prevents the normal maturation of this organelle into an acidic and hydrolytically active compartment, thus avoiding a productive host immune response and activation of the macrophage (Russell, 2001). There is also recent evidence that *M. tuberculosis* is able to rupture its phagosome, thereby it gains access to the host cytosol and induces necrotic cell death of the infected macrophage. Like this, the bacterium escapes innate host defences and favors its own spread to new cells (Simeone *et al.*, 2012; van der Wel *et al.*, 2007).

Approximately 5-10 % of the infected individuals develop active tuberculosis during their life-time. Reactivation of the latent disease is triggered by factors that compromise the immune status of the host, like malnutrition, age, cancer or co-infection with HIV (Ducati *et al.*, 2006). Under such circumstances, the cellular immune system is not able to control the integrity of the granuloma anymore. The granuloma decays into a structureless mass of cell debris and spills viable infectious bacilli into the airways, thus developing a productive infectious cough (Russell, 2007). Besides the pulmonary tuberculosis, about 15% of the patients with active disease develop extra-pulmonary or miliary tuberculosis. Due to the massive growth the bacilli invade into the blood stream and disseminate to various organs of

the host. Liver, spleen, lymph nodes, bones and joints can be affected and beside the typical cough, this inflammatory process produces symptoms like fever, weakness, chest-pain and weight-loss (Ducati *et al.*, 2006).

Mycobacteria

Mycobacteria belong to the family of *Mycobacteriaceae* in the suborder of *Corynebacterineae*, a distinct group of "Gram-positive GC-rich" bacteria within the order of *Actinomycetales*. The genus *Mycobacterium* comprises more than 150 species, most of them being non-pathogenic, saprophytic soil bacteria. However, there are several exceptions which are highly pathogenic to humans, causing tuberculosis (*M. tuberculosis*, *M. bovis*, *M. africanum*), leprosy (*M. leprae*) and Buruli ulcer (*M. ulcerans*) (Cosma *et al.*, 2003; Ducati *et al.*, 2006). Mycobacteria are aerobic-to-facultative anaerobe bacteria with a size of 1-4 µm in length. They are nonmotile, not able to form spores (Traag *et al.*, 2010) nor to produce toxins, with the exception of *M. ulcerans* which produces the toxin mycolactone (George *et al.*, 1999). Mycobacteria weakly stain Gram-positive and can be visualized by Ziehl-Neelsen (acid-fast) staining. The bacilli are rod-shaped and form pale-white, yellow or orange colonies with a rough and wrinkled surface. However, some mycobacteria also have a smooth appearance which is caused by the glycopeptidolipids in their cell envelope.

A rough differentiation of mycobacteria can be made with respect to generation time and pathogenicity. *M. tuberculosis* belongs to the group of pathogenic slow-growing bacteria with a generation time of 20 to 24 hours. In contrast, *M. smegmatis* is a non-pathogenic fast-growing mycobacterium with a generation time of 2 to 3 hours and therefore it is used as model organism in mycobacterial research.

The complete genome sequence of the best characterized and most commonly used laboratory strain, *M. tuberculosis* H37Rv is available. The circular chromosome of 4.4 Mbp consists of approximately 4000 genes and features a high G+C content of 65.6% (Cole *et al.*, 1998). The genome is rich in repetitive DNA sequences, especially insertion elements which are tolerated in intergenic or non-coding regions, but the incidence of natural plasmids has not been observed in *M. tuberculosis* (Ducati *et al.*, 2006). Furthermore, the genome is characterized by a high incidence of genes encoding for enzymes involved in lipid metabolism. This unique feature distinguishes the *M. tuberculosis* genome from any other bacteria. About 250 genes are either involved in lipogenesis for cell envelope synthesis or lipolysis, required for survival inside the host. In *E. coli*, for example, only 50 genes are involved in fatty acid metabolism.

The multitude of genes for lipid metabolism enables mycobacteria to build a complex cell wall structure with a high content of different lipid species (Cole *et al.*, 1998).

The Mycobacterial Cell Envelope

Bacteria are phenotypically classified according to their behaviour in the Gram staining procedure into Gram-positive or Gram-negative bacteria. The difference in staining is due to differences in structure of the cell envelope. The staining dye can be extracted with alcohol from Gram-negative but not from Gram-positive bacteria (Madigan, 2000). The cell envelope of Gram-negative bacteria is rather complex and consists of different layers, the plasma membrane, periplasm, peptidoglycan and the outer membrane. The periplasm is filled with an aqueous solution composed of saccharides, amino acids, peptides and biosynthetic precursors and is only connected to the surrounding medium by porins. The asymmetric, bilayered outer membrane of Gram-negative bacteria resembles the cytoplasmic membrane and consists of lipids, lipopolysaccharides and proteins. Based on the existence of two lipid bilayers Gram-negative bacteria are also termed "diderm" bacteria (Desvaux *et al.*, 2009) (Fig.1). The cell envelope of Gram-positive bacteria is composed of the cytoplasmic membrane, a periplasmic space and a thick layer of peptidoglycan interspersed with proteins, teichoic- and lipoteichoic-acids. Due to the fact that the cytoplasmic membrane is the only bilayer in Gram-positive bacteria, they are also termed "monoderm" bacteria (Desvaux *et al.*, 2009; Seltmann, 2002) (Fig. 1).

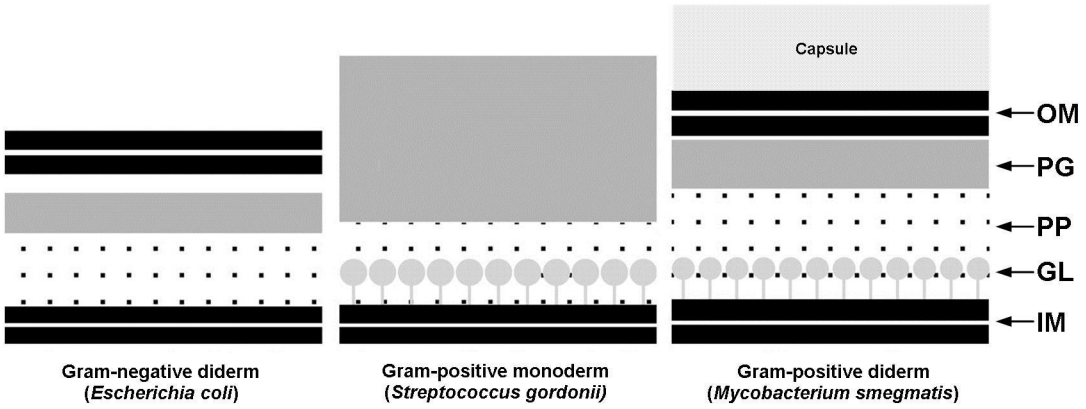


Figure 1: Schematic comparison of the cell wall from Gram-negative diderm, Gram-positive monoderm and Gram-positive diderm bacteria (adapted from (Zuber *et al.*, 2008). All three classes of bacteria have an inner cytoplasmic membrane (IM) followed by a periplasmic space (PP). The granular-layer (GL) is only present in Gram-positive bacteria and mainly composed of penicillin binding proteins (Zuber *et al.*, 2006). The peptidoglycan layer (PG) is much thicker in monoderm bacteria than in diderm bacteria. An outer membrane (OM) is only present in diderm bacteria. Gram-positive diderm bacteria like *M. smegmatis* are further mantled by a capsule.

The mycobacterial cell envelope has a unique and quite complex structure and is highly lipidated, with about 40% of the cell dry mass being lipids, compared to only 10% in *E. coli* (Goren, 1979). It consists of the cytoplasmic membrane, a periplasmic space with adjacent mycolyl-arabinogalactan-peptidoglycan (MAP) complex and noncovalently linked glycans, lipids and proteins which form a capsule-like structure or can be interspersed within the MAP complex (Fig 2). The entirety of structures outside the cytoplasmic membrane is referred to as cell wall.

The cytoplasmic membrane mainly consists of phospholipids, as there are phosphatidylglycerol, phosphatidylethanolamine, phosphatidylinositol and its mannosides. Phosphatidylinositol mannosides (PIMs) are a major component of the cytoplasmic membrane and restricted to bacteria within the order *Actinomycetales*. They also form the base of lipoarabinomannan (LAM), a prominent component of the cell wall (Fig 2).

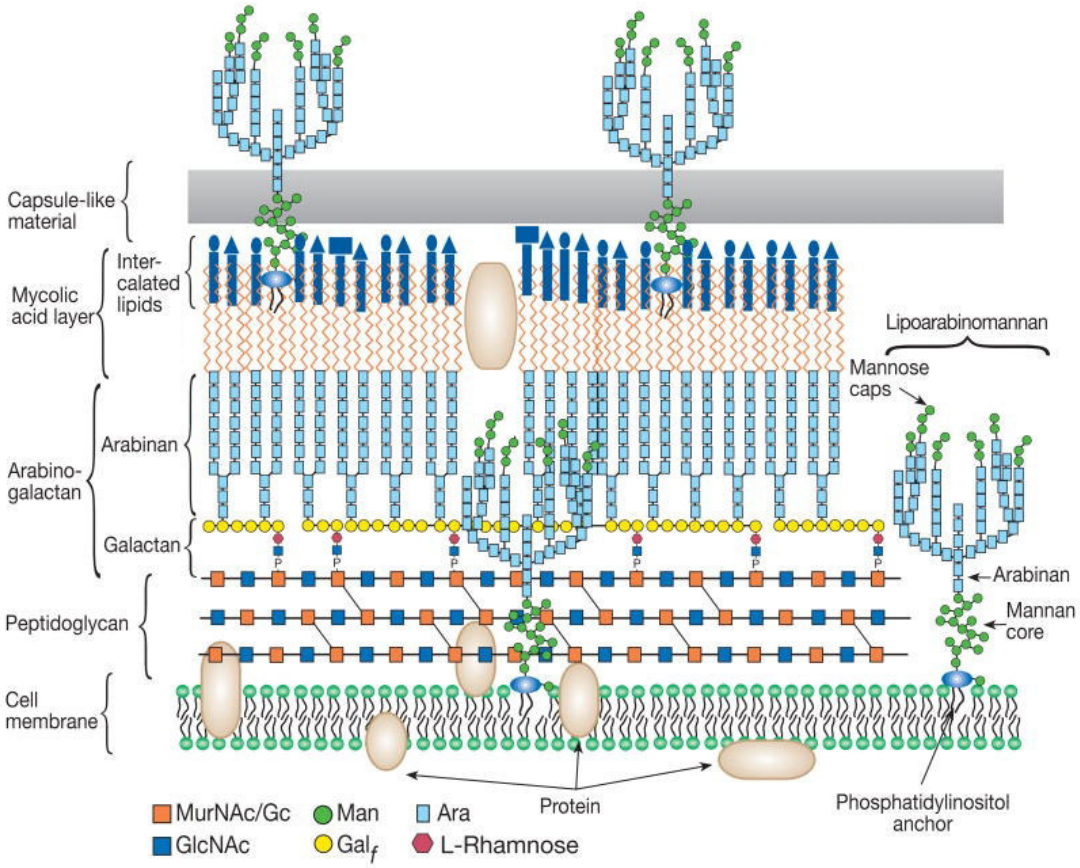


Figure 2: Ultrastructure of the mycobacterial cell envelope (derived from (Brennan & Crick, 2007).

The mycolyl-arabinogalactan-peptidoglycan complex shows some distinct characteristics. Compared to most other bacteria, the peptidoglycan-layer of mycobacteria contains *N*-glycolylmuramic acid instead of *N*-acetylmuramic acid. As another unique feature, the

extensively branched arabinogalactan is phosphodiester-linked to the peptidoglycan and covalently linked to the mycolic acids. Mycolic acids are long-chained α-branched, β-hydroxylated high molecular weight fatty acids. Compared to mycolic acids from other bacteria like corynebacteria, nocardia or rhodococci, mycobacterial mycolic acids are the largest (C70 to C90) and contain double bonds or cyclopropane rings (Brennan & Nikaido, 1995). They are oriented perpendicular to the cytoplasmic membrane forming the outer permeability barrier of the periplasmic space, which is defined at its internal border by the cytoplasmic membrane (Sutcliffe & Harrington, 2004). Free lipids such as phospholipids, trehalose mono- and dimycolates or glycopeptidolipids are intercalating the mycolic acids. The capsule-like layer is composed of proteins, polysaccharides, free lipids such as waxes and glycolipids noncovalently linked to the mycolyl-arabinogalactan-peptidoglycan complex (Brennan & Nikaido, 1995).

Lipoglycans like the most prevalent PIMs and their multiglycosylated analogues lipomannan (LM) and lipoarabinomannan (LAM) are widely interspersed within the mycobacterial cell envelope. As a common feature they all share a conserved phosphatidylinositol anchor and none of them is covalently attached to the cell envelope. LAM are composed of three major structures including the phosphatidylinositol anchor, a D-mannan polymer with D-arabinose side chains and capping motifs at the termini of the arabinose residues.

While PIMs are found in the cytoplasmic membrane among other phospholipids and also in the capsule, the localization of LM and LAM remains unclear. LAMs are supposed to span the whole cell envelope, with their anchor being attached to the cytoplasmic membrane or the mycolic acid layer (Daffe & Draper, 1998). LAM is a heterogenous molecule which differs in terms of capping motifs according to mycobacterial species. Three general classes of LAM have been described: 1. ManLAM with extensive mannose capping as found in *M. tuberculosis* and other slow-growing, pathogenic mycobacteria, 2. PILAM with phosphoinositol capping, found in fast-growing, non-pathogenic mycobacteria like *M. smegmatis* and 3. AraLAM which was found in non-tuberculous *M. chelonae* and remains uncapped (Karakousis *et al.*, 2004).

Although mycobacteria belong to the class of Gram-positive bacteria, it was recently proved that they have a pseudo-outer membrane analogous to the outer membrane of Gram-negative bacteria. The mycolic acid layer constitutes the inner leaflet and the fatty acyl chains of the interspersed lipids intercalating the mycolic acids build up the outer leaflet of this pseudo-

outer membrane (Hoffmann *et al.*, 2008). Based on this, mycobacteria are also termed "diderm" bacteria (Fig. 1).
The pseudo-outer membrane forms a very hydrophobic barrier which together with the low number of pore-forming proteins in the mycobacterial cell wall compared to *E. coli* contribute to the generally low susceptibility to solutes that include many antibiotics and therapeutic agents (Zuber *et al.*, 2008).

The bacterial cell envelope is of major importance in terms of interaction between the bacterium and its environment, notably the host in the case of bacterial pathogens (Sutcliffe & Harrington, 2002). To localize specific proteins to their cell envelopes, almost all bacteria perform post-translational lipid modification to produce membrane-anchored lipoproteins (Sutcliffe & Harrington, 2004). Lipoproteins have been shown to be key players in the interaction of bacterial pathogens with their host.
Lipoproteins are secreted proteins N-terminally modified with fatty acids by the consecutive action of preprolipoprotein diacylglycerol transferase (Lgt), prolipoprotein signal peptidase (LspA) and apolipoprotein *N*-acyltransferase (Lnt). In Gram-negative bacteria, lipoproteins are triacylated carrying a *S*-diacylglyceryl and a *N*-acyl modification (Rezwan *et al.*, 2007a). First direct proof of *N*-acylation of lipoproteins in Gram-positive bacteria was provided by structural analysis of heterologously expressed *M. tuberculosis* lipoprotein LppX in *M. smegmatis*, a fast-growing mycobacterium and Lnt homologues were identified in *M. smegmatis* and *M. tuberculosis* (Tschumi *et al.*, 2009).
In Gram-positive bacteria, cell-associated lipoproteins are anchored in the plasma membrane, whereas in Gram-negative bacteria the vast majority of lipoproteins is released from the plasma membrane and located in the outer membrane. In mycobacteria, lipoproteins localize either in the plasma membrane or in the cell wall (Rezwan *et al.*, 2007b; Sutcliffe & Harrington, 2004). Mycobacterial lipoproteins represent about 2.5% of the *M. tuberculosis* predicted proteome, thus representing an important class of cell envelope proteins that contribute to the virulence of this pathogen. The virulence-related functions of lipoproteins include invasion, signalling, evasion of host defence, and modulation of inflammatory processes. In addition, lipoproteins fulfill functions that contribute to the biology of mycobacteria, like nutrient uptake or cell wall maintenance (Sutcliffe & Harrington, 2004), and thereby indirectly affect host-pathogen interactions. Furthermore, lipoproteins have been shown to be key signalling molecules of both the adaptive and innate immune system and have been identified as major antigens of *M. tuberculosis* (Karakousis *et al.*, 2004).

Lipoproteins are potent agonists of Toll-like receptor 2 (TLR2). Diacylated lipoproteins are recognized by heterodimers of TLR2 and TLR6, while triacylated lipoproteins are recognized by heterodimers of TLR1 and TLR2 with the amide bound fatty acid inserting into a hydrophobic channel of TLR1. The *N*-acyl of the lipoprotein is essential for the heterodimerization of TLR2 and TLR1 and thus the initiation of TLR1/2 signalling (Jin *et al.*, 2007; Kang *et al.*, 2009).

TLR-2 signalling has an important pleiotropic, but seemingly antagonistic role in immune responses to *M. tuberculosis*. TLR-2 signalling enhances both innate and adaptive immune responses, but it is also involved in downregulation of some immune functions. *M. tuberculosis* 19 kDa lipoprotein LpqH, Rv3763, for example is identified as an antigen and TLR-2 agonist which promotes TLR-2 dependent bactericidal responses and innate immune functions in macrophages early in infection. But, during later stages prolonged TLR-2 signalling inhibits MHC class II expression and antigen processing and presentation, thereby decreasing recognition by T cells (Harding & Boom, 2010). In addition, LpqH is as a cell wall located adhesin enhancing the phagocytosis of the bacterium by macrophages. The negative-feedback regulation in macrophages seems to help *M. tuberculosis* to evade the host immune response and persist as long-term infection (Diaz-Silvestre *et al.*, 2005).

Aim of the study

Recent studies demonstrated an important role of lipoproteins and their synthesizing enzymes in the biology of mycobacteria (Rampini *et al.*, 2008). Due to their involvement in virulence, they represent potential drug targets or candidates for novel vaccines. The modification of lipoproteins with fatty acids seems to play a crucial role for lipoprotein function, its localization and interaction with TLRs. Recently, the molecular structure of *M. tuberculosis* lipoprotein LppX heterologously expressed in *M. smegmatis* was determined and provided first direct proof of *N*-acylation of a mycobacterial model lipoprotein. However, the exact molecular structure of the membrane anchor so far remains poorly defined.

The goal of this study was, to elucidate if *N*-acylation represents a general modification for mycobacterial lipoproteins and lipoproteins in *Streptomyces ssp.*. *Streptomyces ssp*, like mycobacteria, are GC-rich Gram-positive bacteria and belong to the order of *Actinomycetales*. But, unlike their relatives from the subfamily of *Corynebacterineae* they are monoderm bacteria, lacking an outer membrane-like structure.

Aim of the study:

- Characterize the structure of the membrane anchor of lipoproteins in fast-growing *M. smegmatis*
- Characterize the structure of the membrane anchor of lipoproteins in slow-growing *M. bovis* BCG
- Identify genes responsible for *N*-acylation in slow-growing mycobacteria
- Characterize the structure of the membrane anchor of lipoproteins and genes involved in lipoprotein synthesis in *Streptomyces scabies*

CHAPTER 1

BACTERIAL LIPOPROTEINS: BIOGENESIS, VIRULENCE/PATHOGENICITY AND TRAFFICKING

Hajime Tokuda, Peter Sander, Bok Luel Lee, Suguru Okuda, Thomas Grau, Andreas Tschumi, Juliane K. Brülle, Kenji Kurokawa and Hiroshi Nakayama

Chapter 4 in: Bacterial Membranes: Structural and Molecular Biology. (2014) p.133-177. Remault, H., Fronzes, R. (eds). Caister Academic Press, Norfolk, UK

ABSTRACT

The mechanisms underlying the biogenesis and outer membrane sorting of lipoproteins have been mostly clarified in *Escherichia coli*. Three enzymes catalyse the post-translational modification of lipoproteins with a membrane anchor comprising a thioether-linked diacylglycerol and an amide-linked fatty acid. The Lol system, comprising five Lol proteins, mediates the sorting of lipoproteins to the outer membrane. The three enzymes and five proteins are essential for *E. coli* and are widely conserved in Gram-negative diderms having cytoplasmic and outer membranes. High G+C content Gram-positive bacteria such as those belonging to the genera *Mycobacterium* and *Corynebacterium* have an outer membrane-like structure. The structure and biogenesis of the cell envelope have been intensively studied in mycobacterial species including medically important *Mycobacterium tuberculosis*. Mycobacterial lipoproteins are triacylated, like those of Gram-negative diderms. The enzyme catalysing *N*-acylation was recently identified. The lipoprotein biosynthesis pathway is important for virulence of *M. tuberculosis*. The functions of individual mycobacterial lipoproteins are discussed in relation to envelope biogenesis, virulence and influence on immune systems. The N-terminal structures of lipoproteins of low G+C content Gram-positive monoderms are surprisingly diverse. An enzyme catalysing the *N*-acylation of the N-terminal Cys of lipoproteins has not been found in this class of bacteria. However, lipoproteins of *Staphylococcus* species are *N*-acylated. Moreover, three novel structures of lipidated N-terminal Cys were revealed on mass spectrometric analyses. A possible biosynthetic pathway generating these structures is discussed.

INTRODUCTION

Bacterial lipoproteins constitute a class of membrane proteins functioning outside of the cytoplasmic membrane. After lipoprotein precursors have been translocated across the cytoplasmic membrane, the conserved N-terminal Cys residue of lipoproteins is modified with lipids, which anchor the lipoproteins to membranes. The functions of lipoproteins are diverse; they are involved in drug export, substrate import, construction of extracytoplasmic structures, molecular chaperone activity in the extracytoplasm, activation of Toll-like receptor (TLR) of host cells, etc.

Gram staining is an empirical method for differentiating bacterial species into two large groups, Gram-positive and Gram-negative bacteria. It is almost always the first step in the identification of bacterial organisms, particularly in the medical microbiology laboratory. The Gram-reactivity of a bacterium is based on the ability of the cell to retain the purple primary staining agent crystal violet after undergoing Gram staining. This ability is not based on the actual structure of the cell envelope of the bacterium, although there is often a correlation. It was therefore proposed that bacteria should be classified whether they have one (monoderm) or two (diderm) membranes (Desvaux *et al.*, 2009; Gupta, 1998). Gram-negative bacteria are mostly diderms with a cytoplasmic membrane, a thin peptidoglycan layer and an outer membrane that acts as a permeability barrier to small hydrophilic molecules (Fig. 1). Typical members are Proteobacteria, e.g. *Escherichia coli*. High GC Gram-positive bacteria are generally diderms while low GC ones are monoderms. Low GC Gram positive bacteria are covered by multiple peptidoglycan layers (Fig. 1).

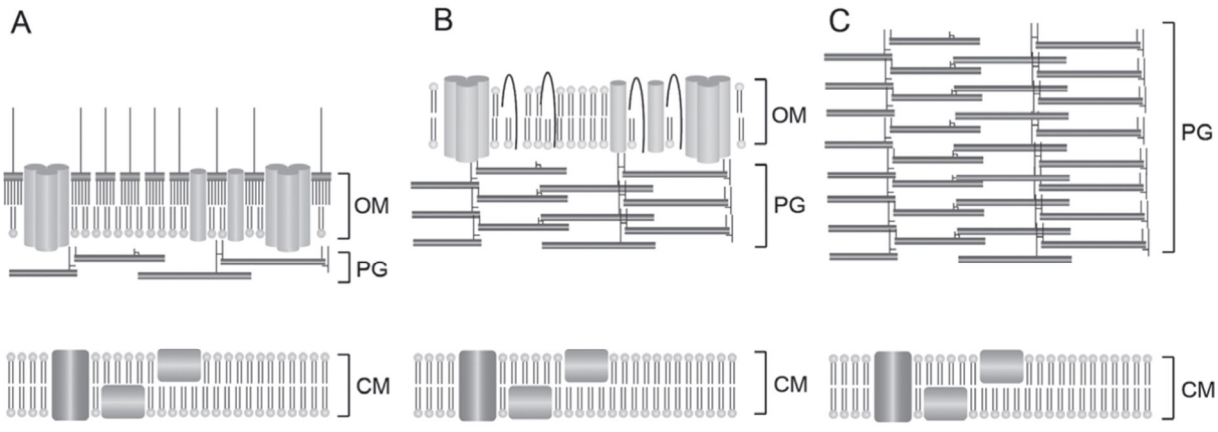


Fig. 1 Schematic representation of the envelope structures of three bacteria, *E. coli* (A), *M. tuberculosis* (B), and *S. aureus* (C), discussed mainly in this chapter. CM, PG and OM represent the cytoplasmic membrane, peptidoglycan and outer membrane, respectively.

Lipoproteins in monoderm bacteria are only anchored to the outer surface of the cytoplasmic membrane. On the other hand, ones in a typical diderm, *Escherichia coli*, are anchored to the periplasmic surface of both the cytoplasmic (inner) and outer membranes. Lipoproteins anchored to the outer surface of the outer membrane are also known in Gram-negative diderms, although such lipoproteins are few in *E. coli*. The localization of lipoproteins in Gram-positive diderms remains largely unknown.

The pathways generating mature lipoproteins from their precursors have been extensively studied in *E. coli*. It has been established that lipoproteins in Gram-negative diderms are triacylated through three sequential modification reactions (Fig. 2). The first enzyme, Lgt,

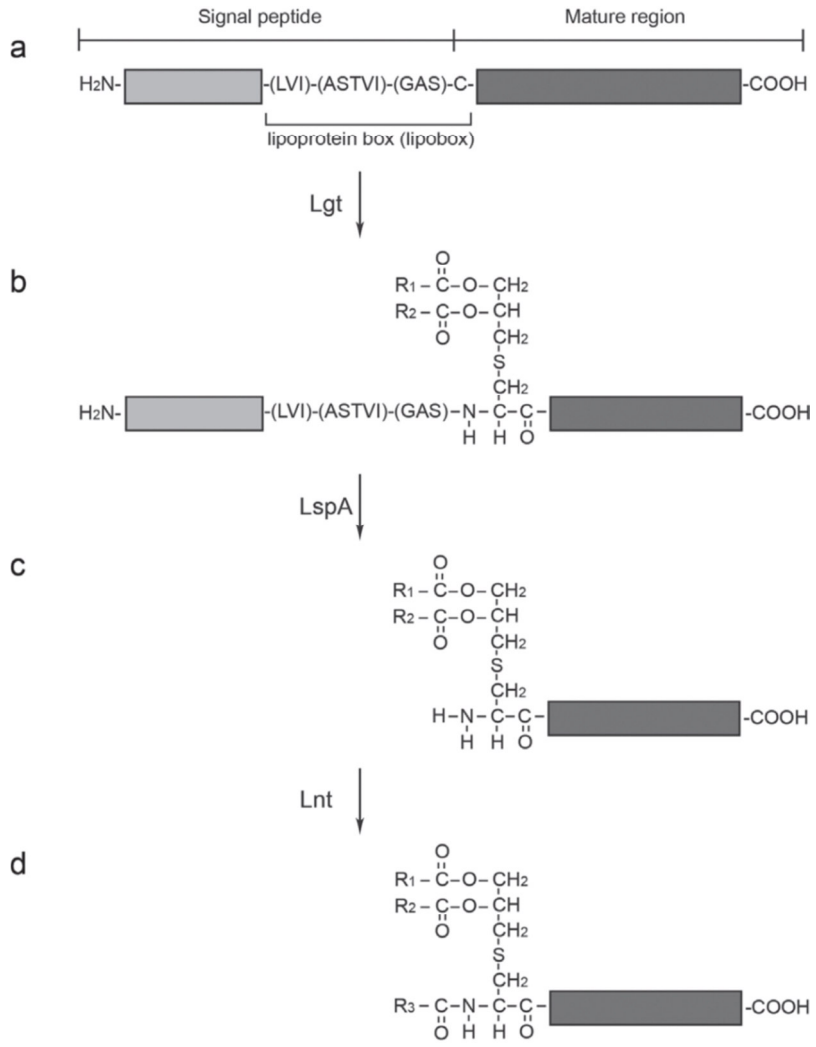


Fig. 2 Lipoprotein processing in bacteria. The lipoprotein maturation pathway found in Gram-negative diderm is shown. Essentially the same pathway is present in Gram-positive diderms. A lipoprotein precursor [pre-prolipoprotein, (a)] is modified by phosphatidylglycerol/ prolipoprotein diacylglyceryl transferase (Lgt), a thioether linkage being formed between the N-terminal Cys of the mature region and diacylglycerol. The signal sequence of diacylated prolipoprotein (b) is cleaved by prolipoprotein signal peptidase (LspA). The apolipoprotein (c) is *N*-acylated by phospholipid/apolipoprotein transacylase (Lnt), and becomes a triacylated mature lipoprotein (d). An Lnt homologue is not found in low GC content Gram-positive bacteria, although some of lipoproteins are triacylated.

modifies the universally conserved Cys in the so-called 'lipoprotein box or lipobox' with a diacylglycerol moiety derived from phosphatidylglycerol (Sankaran & Wu, 1994) when lipoprotein precursors are translocated from the cytoplasm to the periplasmic side of the cytoplasmic membrane. The lipobox is a consensus sequence, (LVI)-(ASTVI)-(GAS)-C, comprising the C-terminal three residues of the signal peptide and the N-terminal Cys in the mature region (http://www.mrc-lmb.cam.ac.uk/genomes/dolop/). The second enzyme, LspA or signal peptidase II, cleaves the signal peptide of the diacylated lipoprotein precursors, rendering the Cys residue a new N-terminus. The last enzyme, Lnt, catalyses the aminoacylation of the N-terminal Cys. It has been thought that triacylated lipoproteins are limited to Gram-negative bacteria because Lnt homologues have not been found in Gram-positive bacteria. However, Lnt homologues were recently found in a Gram-positive diderm (Tschumi *et al.*, 2009). Moreover, it has been reported that the amino group of Cys are also diversely modified even in Gram-positive monoderms (Asanuma *et al.*, 2011; Kurokawa *et al.*, 2009; Kurokawa *et al.*, 2012a; Nakayama *et al.*, 2012).

The aims of this chapter are to discuss, based on recent findings, the mode of N-terminal modification, the enzymes involved in lipoprotein processing, the physiological functions of lipoproteins, and the sorting of lipoproteins in both monoderms and diderms. The Lyme disease spirochaete *Borrelia burgdorferi* is a Gram-negative diderm that possesses many lipoproteins (Fraser *et al.*, 1997), but has no lipopolysaccharides (LPS). Readers interested in spirochaete lipoproteins are referred to reports by Haake (Haake, 2000) and Schulze *et al.* (Schulze *et al.*, 2010).

LIPOPROTEINS IN GRAM-NEGATIVE DIDERM BACTERIA

The outer membrane is an essential organelle for most Gram-negative diderms, and comprises phospholipids, lipopolysaccharides (LPS), membrane spanning β-barrel proteins, and lipoproteins. These components are transported from the cytoplasmic to the outer membrane through the periplasm. Recent studies revealed that many lipoproteins are involved in the transport of and insertion of these components into the outer membrane. Among these four components, the sorting of lipoprotein to the outer membrane has been most extensively studied and clarified.

Lipoprotein-processing enzymes

Three well-conserved enzymes sequentially mediate the maturation reactions (Fig. 2) (Sankaran & Wu, 1994):

1 Phosphatidylglycerol/prolipoprotein diacylglyceryl transferase (Lgt) catalyses the transfer of the diacylglyceryl moiety from phosphatidylglycerol (PG) in the cytoplasmic membrane to the N-terminal Cys residue in the mature region through a thioether linkage.
2 Lipoprotein-specific signal peptidase (LspA or signal peptidase II) then cleaves the hydrophobic signal peptide.
3 The third enzyme, phospholipid/apolipoprotein transacylase (Lnt), attaches an acyl chain to the free amino group of the N-terminal Cys residue.

These three enzymes are essential for the growth of *E. coli* and *Salmonella enterica* serovar Typhimurium (Wu, 1966). The *N*-acylation of lipoproteins by Lnt has been shown to be critically important for their sorting *in vitro* (Fukuda *et al.*, 2002). An *E. coli* mutant, in which the chromosomal *lnt* gene was placed under an arabinose promoter, revealed *in vivo* that apolipoproteins were accumulated in the cytoplasmic membrane in the absence of arabinose (Robichon *et al.*, 2005). Lnt was found to have at least six transmembrane helices with a nitrilase-like periplasmic domain containing the catalytic triad Glu267–Lys335–Cys387 (Robichon *et al.*, 2005; Vidal-Ingigliardi *et al.*, 2007). It was proposed for the mechanism of the third lipid attachment that Lnt attacks the *sn*-1 acyl chain of a phospholipid and thereby forms a thioester acyl-enzyme intermediate, followed by the transfer of the acyl group to the α-amino group of the diacylglyceryl Cys (Buddelmeijer & Young, 2010). Recently, the *N*-acyltransferase activity of *E. coli* Lnt was found to strongly depend on phospholipid head group and acyl chain composition (Hillmann *et al.*, 2011). Lnt prefers phospholipids carrying a small head group, such as phosphatidylethanolamine (PE) or PG, as an acyl donor, and exhibits significant activity only when an unbranched acyl chain is present at position *sn*-1 with an unsaturated one at position *sn*-2 (Hillmann *et al.*, 2011). Finally, the *lnt* gene was found to be dispensable when an ABC transporter, LolCDE complex, is overproduced in *E. coli* cells lacking the major outer membrane lipoprotein Lpp (Narita & Tokuda, 2011).

The Lol system

In Gram-negative bacteria such as *E. coli*, triacylated mature lipoproteins are transported to the outer membrane by the Lol (localization of lipoproteins) system or are retained in the cytoplasmic membrane, depending on the residue at position 2, which is located adjacent to the triacylated N-terminal Cys. The Lol system is composed of five proteins, periplasmic carrier protein LolA, outer membrane lipoprotein receptor LolB, and cytoplasmic membrane ABC transporter LolCDE (Tokuda & Matsuyama, 2004). These five Lol proteins are conserved in Gram-negative bacteria, and are essential for the growth of *E. coli* and *S. enterica*, as the three processing enzymes are. Membrane subunits LolC and LolE of the LolCDE complex are only conserved in diderm bacteria. The conservation of LolB is much lower, and seems to be limited to β- and γ-proteobacteria among diderms. It is interesting that the proteins involved in the early step of lipoprotein processing or sorting are more strongly conserved than those involved in a later step (Okuda & Tokuda, 2011). The diderm bacteria, in which LolB is not found, might have some other outer membrane proteins or lipoproteins that can compensate for the LolB function.

Periplasmic carrier protein LolA

The finding of a lipoprotein-specific carrier protein, LolA, revealed how hydrophobic lipoproteins cross the hydrophilic periplasmic space and how the membrane specificity of lipoproteins is determined (Matsuyama *et al.*, 1995). Spheroplasts secrete various periplasmic and hydrophobic β-barrel outer membrane proteins into the medium, whereas outer membrane-specific lipoproteins, such as Lpp, are retained in the cytoplasmic membrane of spheroplasts. However, Lpp is secreted into the spheroplast medium when the periplasmic fraction is added externally. The Lpp-releasing activity in the periplasmic fraction has been purified, and a water-soluble protein named LolA has been identified. LolA also releases other outer membrane-specific lipoproteins such as Pal, BamC, Slp and RlpA, whereas cytoplasmic membrane-specific lipoproteins AcrA and NlpA remain in spheroplasts even when LolA is added. These observations suggest that the localization of lipoproteins in the outer membrane is determined by whether or not they are released from the cytoplasmic membrane. Lipoproteins released from spheroplasts were found to form a water-soluble complex with LolA in a stoichiometry of 1:1. Moreover, when LolA was genetically depleted, outer membrane-specific lipoproteins were accumulated in the cytoplasmic membrane

(Tajima *et al.*, 1998). LolA was thus found to play essential roles in the release of lipoproteins from the cytoplasmic membrane and their transport through the hydrophilic periplasm.

Outer membrane receptor LolB

The LolA–lipoprotein complex described above has been isolated from a spheroplast supernatant, and then incubated with cytoplasmic and outer membranes. It was found that lipoproteins were specifically transferred to the outer membrane, but not to the cytoplasmic membrane (Matsuyama *et al.*, 1995). This observation suggested that an unidentified factor in the outer membrane accepts lipoproteins from LolA and then mediates their outer membrane localization. Solubilized outer membrane proteins have been fractionated and then reconstituted into proteoliposomes to identify this factor. The outer membrane lipoprotein named LolB was thus identified as a lipoprotein receptor anchored to the outer membrane (Matsuyama *et al.*, 1997). LolB depletion caused the accumulation of outer membrane-specific lipoproteins both in the periplasm as a complex with LolA and in the cytoplasmic membrane (Tanaka *et al.*, 2001). A LolB derivative, mLolB, lacking an N-terminal acyl chain is able to form a complex with lipoproteins as LolA does, indicating that LolA and LolB are functionally similar although their amino acid sequences are dissimilar. The modes of lipoprotein binding to LolA and LolB, and the importance of the lipid anchor of LolB are discussed in a later section.

Cytoplasmic membrane ABC transporter LolCDE

The outer membrane is the final destination of lipoproteins, and they are never released into the periplasm even in the presence of LolA. The release of outer membrane-specific lipoproteins from right-side-out membrane vesicles occurs in ATP- and LolA-dependent manners (Yakushi *et al.*, 1998). These observations indicate that there is an ATPase that catalyses the release of lipoproteins exclusively in the cytoplasmic membrane. Cytoplasmic membrane proteins have been solubilized, fractionated and reconstituted into liposomes together with outer membrane-specific lipoproteins to examine the lipoprotein-releasing activity of each fraction in the presence of ATP and LolA. The LolCDE complex, which belongs to the ABC (ATP-binding cassette) transporter super family, was thus identified (Yakushi *et al.*, 2000). This complex consists of LolC, LolD, and LolE subunits in a stoichiometry of 1:2:1. LolC and LolE are membrane subunits, each of which has four

membrane-spanning regions and one large periplasmic loop (Yasuda *et al.*, 2009). LolD is a nucleotide-binding subunit with conserved ABC signature, Walker A, and Walker B motifs.
ABC transporters generally mediate the ATP-dependent transport of diverse substrates across the lipid bilayers of various organisms. They are composed of four domains; two transmembrane domains (TMDs), and two nucleotide-binding domains (NBDs). ABC transporters in eukaryotic cells mainly function as drug exporters, of which the TMDs and NBDs are expressed in a single polypeptide chain. On the other hand, most bacterial ABC transporters are involved in the uptake of nutrients with a specific substrate-binding protein, and have four domains frequently in separate polypeptide chains (Davidson & Chen, 2004; Dawson *et al.*, 2007). Some bacterial ABC transporters mediate the export of various drugs, inhibitors, polysaccharides and proteins.
LolCDE, having eight membrane-spanning regions in total, is assumed to be a derivative of a bacterial exporter, whereas exporters generally have at least twelve membrane-spanning regions. MacB, another bacterial ABC transporter, also has eight transmembrane regions in total, and exports enterotoxin II from the periplasm to the exterior in collaboration with a membrane fusion protein, MacA, and a multifunctional outer membrane channel, TolC (Kobayashi *et al.*, 2003; Yamanaka *et al.*, 2008). MacB is predicted to have a large periplasmic loop like LolC and LolE have. Both LolCDE and MacB release substrates after they have been translocated across the cytoplasmic membrane. Their characteristic structures might be related to their unique functions.

Sorting of lipoproteins to the outer membrane of Gram-negative bacteria

The molecular events occurring during the transport of lipoproteins from the cytoplasmic to the outer membrane through the periplasm have been extensively studied in *E. coli* (Fig. 3). Outer membrane-specific lipoproteins are recognized by LolCDE in the cytoplasmic membrane and then transferred to LolA in an ATP-dependent manner. Lipoproteins cross the periplasmic space as a complex with LolA and are transferred to LolB anchored to the outer membrane, followed by incorporation of the lipid moiety of the lipoproteins into the outer membrane. In this section, the mechanisms underlying the lipoprotein transfer reactions through five Lol proteins are discussed.

Sorting signals of lipoproteins

Yamaguchi *et al.* (Yamaguchi *et al.*, 1988) first indicated the importance of Asp at position 2 for the cytoplasmic membrane localization of lipoproteins in *E. coli* by demonstrating that the substitution of Ser with Asp at position 2 caused the retention of an outer membrane-specific lipoprotein in the cytoplasmic membrane. Moreover, substitution of Asp at position 2 of a cytoplasmic membrane-specific lipoprotein with another residue was shown to cause the outer

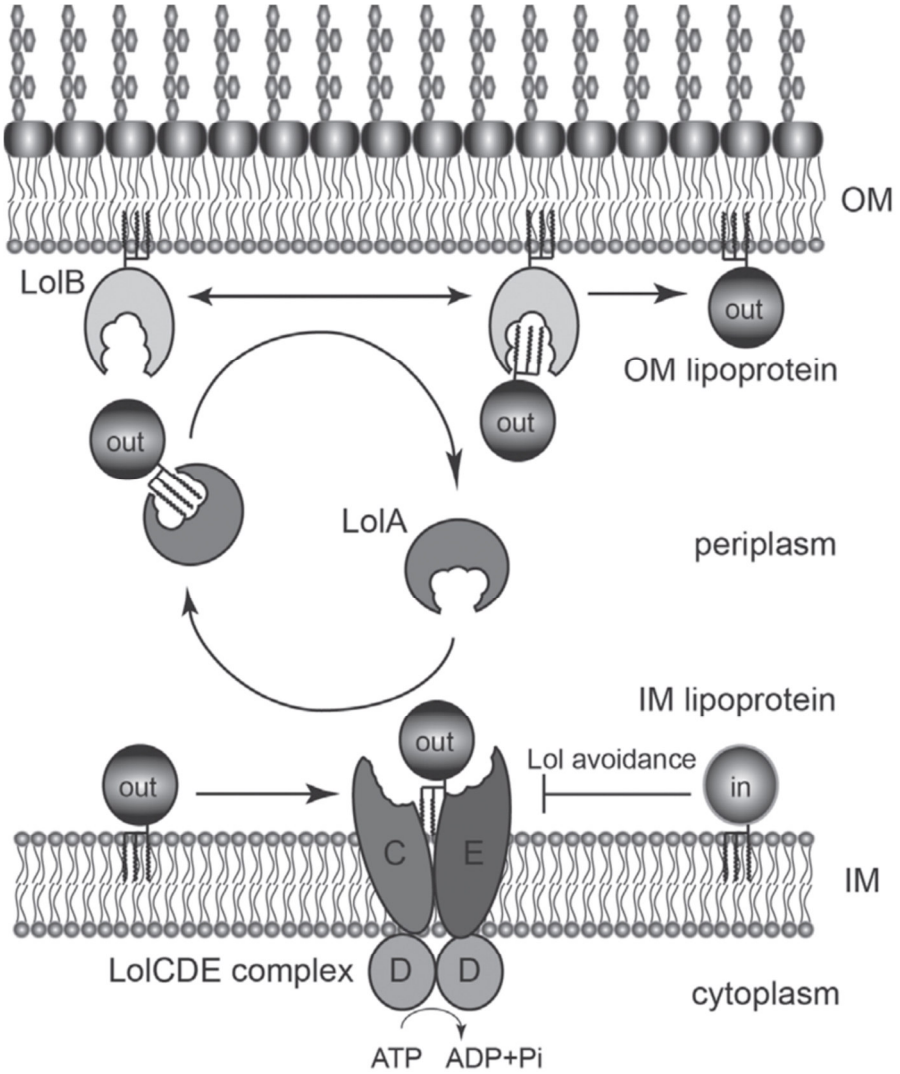


Fig. 3 The Lol pathway. An ABC transporter, the LolCDE complex, recognizes an outer membrane-specific mature lipoprotein in the cytoplasmic membrane. The Asp residue at position 2 functions as an 'Lol avoidance signal' and prevents the recognition of lipoproteins by LolCDE. This causes the retention of lipoproteins in the cytoplasmic membrane. The released lipoprotein forms a complex with a periplasmic chaperone, LolA, in an ATP-dependent manner. The LolA–lipoprotein complex crosses the periplasm and reaches the outer membrane, where the lipoprotein is transferred from LolA to LolB. LolB is itself a lipoprotein anchored to the outer membrane and mediates the incorporation of lipoproteins into the periplasmic leaflet of the outer membrane.

membrane-localization of the lipoprotein. It was reported later that the residue at position 3 variously affects the Asp-dependent cytoplasmic membrane localization of lipoproteins (Gennity & Inouye, 1991). To examine this '+2 rule', the residues at positions 2 and 3 were systematically changed and the LolA-dependent release of lipoproteins from the cytoplasmic membrane was examined. It was found that only Asp at position 2 caused the retention of a lipoprotein in the cytoplasmic membrane when the residue at position 3 was Ser (Terada *et al.*, 2001). The residue at position 3 was then found to variously affect the cytoplasmic membrane retention of lipoproteins by Asp at position 2. Asp at position 2 was the strongest cytoplasmic membrane signal when Asp, Glu or Gln was at position 3 (Terada *et al.*, 2001). The combination of Asp and Asn or Asp and Arg at positions 2 and 3, respectively, also functioned as a relatively strong cytoplasmic membrane signal. Native cytoplasmic membrane-specific lipoproteins of *E. coli* have Asp at position 2 and Asp, Glu or Gln at position 3. Taken together, these results suggested that the negative charge or amide group at position 3 increases the potency of Asp at position 2 as a cytoplasmic membrane retention signal. Asn at position 2 exceptionally functions as a cytoplasmic membrane retention signal only when the residue at position 3 is Asp, as in the case of *E. coli* native lipoprotein AcrE (Klein *et al.*, 1991; Seiffer *et al.*, 1993). It remains to be elucidated why Lys or His with an amide group at position 3 decreases the retention of cytoplasmic membrane-specific lipoproteins (Terada *et al.*, 2001).

The bacterial genome carries a number of putative lipoprotein genes (Babu *et al.*, 2006; Juncker *et al.*, 2003). In the family Enterobacteriaceae, such as *S. enterica* serovar Typhimurium, *Shigella flexneri*, *Yersinia pseudotuberculosis*, *Erwinia carotovora*, and *Klebsiella oxytoca*, the '+2 rule' seems to be conserved, as judged on the direct visualization of fluorescence-labelled lipoproteins *in vivo* (Lewenza *et al.*, 2006). However, the sorting signal may vary in other bacteria even though they possess Lol protein homologues. For example, the residue at position 2 of MexA, a cytoplasmic membrane lipoprotein constituting the multidrug efflux pump of *Pseudomonas aeruginosa*, is Gly. Chimeric lipoproteins consisting of various regions of MexA and outer membrane-specific lipoprotein OprM revealed that Lys and Ser at positions 3 and 4, respectively, but not Gly at position 2, are the typical cytoplasmic membrane signals of *P. aeruginosa*. Although Asp at position 2 also caused the cytoplasmic membrane retention of lipoproteins, the residues at positions 3 and 4 were found to be the innate sorting signals for lipoproteins in *P. aeruginosa (Narita & Tokuda, 2007)*. Indeed, the Lys-Ser signal functioned as the cytoplasmic membrane retention signal in proteoliposomes reconstituted with a LolCDE homologue purified from

P. aeruginosa, while reconstitution of *E. coli* LolCDE caused the release of this lipoprotein. It should be noted that LolA purified from not only *E. coli* but also *P. aeruginosa* functions in this release from proteoliposomes (Tanaka *et al.*, 2007). Asp at position 2 functions as a cytoplasmic membrane retention signal whether LolCDE is derived from *E. coli* or *P. aeruginosa*. These findings suggest that the mechanisms underlying the Lol-dependent localization of lipoproteins are similar in *E. coli* and *P. aeruginosa*, although the cytoplasmic membrane retention signal of lipoproteins depends on the properties of LolCDE. The lipoprotein-sorting signals in Gram-negative bacteria seem to be more diverse than previously expected.

It was once thought to be possible that Asp at position 2 inhibits *N*-acylation and therefore causes retention of lipoproteins in the cytoplasmic membrane. However, this was found not to be the case, and cytoplasmic membrane-specific lipoproteins with Asp at position 2 were shown to be *N*-acylated (Fukuda *et al.*, 2002). It was also revealed that non-*N*-acylated lipoproteins (apolipoproteins) are hardly released from the cytoplasmic membrane, indicating that LolCDE recognizes triacylated lipoproteins. As mentioned above, the *lnt* gene can be deleted when cells overproduce LolCDE and lack either Lpp or the genes encoding the three transpeptidases, which form cross-links between Lpp and peptidoglycan (Narita & Tokuda, 2011). The diacylated lipoproteins in this *Δlnt* mutant were poor substrates for LolCDE, therefore overproduction of LolCDE was required for their correct sorting.

The Lol avoidance signal

The reason why Asp at position 2 causes the retention of lipoproteins in the cytoplasmic membrane has been studied by means of reconstitution experiments. Pal(S2D), a derivative of the outer membrane-specific lipoprotein Pal possessing Asp in place of Ser at position 2, was not released from the proteoliposomes reconstituted with LolCDE in the presence of LolA and ATP (Yakushi *et al.*, 2000). The ATPase activity of LolCDE was stimulated by Pal, but not by Pal(S2D) (Masuda *et al.*, 2002). Moreover, the release of an outer membrane-specific lipoprotein from proteoliposomes was completely inhibited by an excess amount of Pal in the same proteoliposomes, whereas an excess amount of Pal(S2D) had no effect (Masuda *et al.*, 2002). These observations suggested that Asp at position 2 functions as a 'Lol avoidance signal' and thereby prevents the recognition of lipoproteins by LolCDE. Indeed, various LolCDE mutants have been isolated that localize cytoplasmic membrane-specific lipoproteins to the outer membrane (Narita *et al.*, 2003; Sakamoto *et al.*, 2010). These findings indicate

that both LolA and LolB recognize lipoproteins even though they have Asp at position 2 when the LolCDE mutants release those lipoproteins. An altered conformation of LolCDE most likely accounts for the suppression of the Lol avoidance function of Asp at position 2.

The mechanism underlying lipoprotein-recognition by LolCDE was further examined with proteoliposomes reconstituted with LolCDE and chemically modified lipoproteins (Hara *et al.*, 2003). A Pal derivative with Cys in place of Ser at position 2 was released from reconstituted proteoliposomes even after the Cys residue had been modified with SH-specific reagents. Since *N*-acylation is essential for recognition by LolCDE, it seemed likely that LolCDE recognizes the N-terminal Cys residue modified with three acyl chains, the sole common structure of lipoproteins. The oxidation of Cys to cysteic acid resulted in the generation of a Lol avoidance signal. The calculated distances between Cα and the negative charges of Asp and cysteic acid are very similar. On the other hand, Glu at position 2 does not function as a Lol-avoidance signal although it is negatively charged and has similar properties to Asp. These results suggest that the distance between the negative charge and Cα at position 2 is critical for the LolCDE avoidance function.

E. coli membranes contain 70–75% PE, 20–25% PG, and about 5% cardiolipin (CL). Examination of the release of lipoproteins from proteoliposomes reconstituted with various phospholipids revealed that Asp at position 2 functioned as the Lol avoidance signal in proteoliposomes reconstituted with phosphatidylcholine (PC) possessing a positive charge. In contrast, the Lol avoidance function of Asp was abolished if the proteoliposomes were reconstituted with *E. coli* phospholipids pretreated with an amine-specific reagent (Hara *et al.*, 2003). These observations suggested that the steric and electrostatic complementarity between Asp at position 2 and phospholipids such as PE having a positive charge is important for the Lol avoidance mechanism. However, an *E. coli* mutant unable to produce PE can grow on a medium supplemented with a high concentration of magnesium (DeChavigny *et al.*, 1991), and the sorting of lipoproteins was found to be normal in the mutant cells (Miyamoto & Tokuda, 2007), indicating that the positive charge of phospholipids is not essential for the Lol avoidance signal. The phospholipids in this mutant were comprised 50% CL and 50% PG. CL is thought to exhibit a non-bilayer phospholipid property in the presence of a high concentration of magnesium (Rietveld *et al.*, 1995). Both the release of lipoproteins and the ATPase activity of LolCDE were stimulated by magnesium when proteoliposomes were reconstituted with CL alone (Miyamoto & Tokuda, 2007). However, the lipoproteins with Asp at position 2 were also released under these conditions. In contrast, PG added to CL-liposomes increasingly suppressed the release of lipoproteins with Asp and lowered the

ATPase activity of LolCDE. Taken together, these findings indicate that phospholipids have diverse effects on lipoprotein sorting in *E. coli*. It remains unknown how PG negatively affects the release of lipoproteins with Asp at position 2.

Isolation of liganded LolCDE

As far as reported, no ABC transporter has been co-purified with its substrate. On the other hand, a liganded form of LolCDE could be purified if ATP was absent during its purification. Various lipoproteins co-purified with LolCDE were all outer membrane-specific, but not cytoplasmic membrane-specific (Ito *et al.*, 2006), indicating that the liganded LolCDE represents an intermediate of the lipoprotein release reaction in the cytoplasmic membrane. LolCDE liganded exclusively with Pal in a molar ratio of 1:1 was then purified. Taking advantage of this novel liganded LolCDE, the molecular events involved in the release of lipoproteins by LolCDE was divided into the following steps:

1 Outer membrane-specific lipoproteins are recognized by LolCDE in an ATP-independent manner unless the lipoprotein has Asp at position 2. The affinity of LolD for ATP increases upon lipoprotein binding.
2 ATP-binding to LolD decreases the strength of the hydrophobic interaction between lipoproteins and LolCDE through a conformational change of LolCDE.
3 Lipoproteins are transferred from LolCDE to LolA upon the hydrolysis of ATP. However, when LolA is not present, lipoproteins remain associated with LolCDE even if ATP is hydrolysed. This step involves a conformational change of LolA, as discussed later.

The crystal structure of the LolCDE complex has not been solved while a methanococcal LolD homologue exhibits a very similar tertiary fold to those of the ATPase subunits of other ABC transporters (Smith *et al.*, 2002; Yuan *et al.*, 2001). Since LolCDE can be assumed to be a variant of ABC exporter, the molecular mechanism of LolCDE might be similar to that proposed for crystallized ABC exporter Sav1866 of *Staphylococcus aureus* (Dawson & Locher, 2006; Dawson & Locher, 2007).

Structures of LolA and LolB

Both LolA and LolB transiently bind lipoproteins and then transfer the bound lipoproteins to LolB and the outer membrane, respectively. Their functions are similar but their amino acid sequences exhibit no similarity. In contrast, the crystal structures of *E. coli* LolA and LolB solved at 1.65 and 1.9 Å resolution, respectively, exhibit remarkable similarity to each other (Takeda *et al.*, 2003). Both LolA and LolB form an incomplete β-barrel structure composed of 11 anti-parallel β-strands with a lid comprising three α-helices (Fig. 4). The incomplete β-barrel and the lid form a hydrophobic cavity, which was shown to bind the acyl chains of lipoproteins, as discussed later. Their overall structures are similar, but two differences are

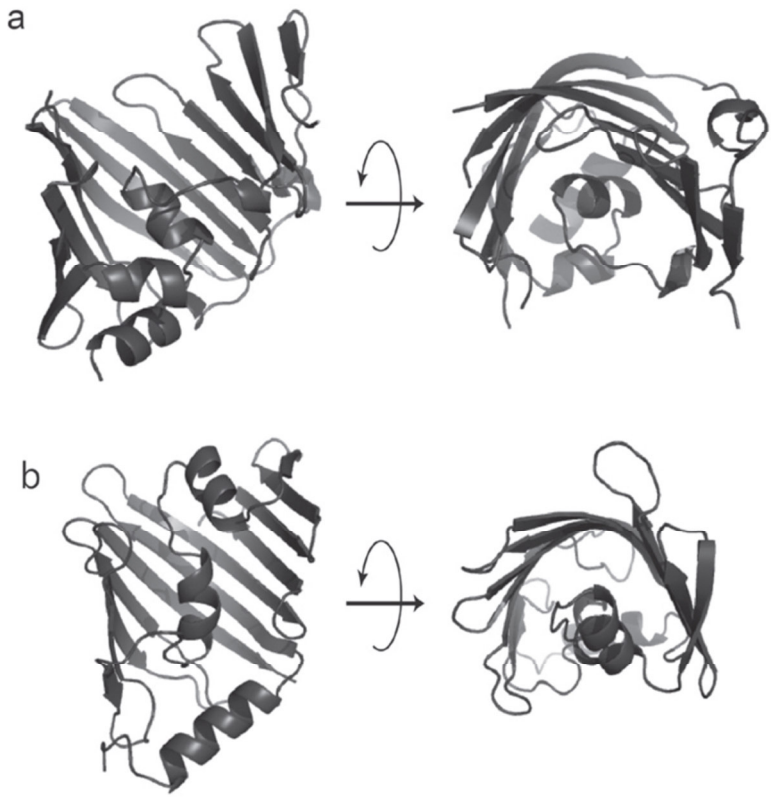


Fig. 4 Crystal structures of (a) LolA and (b) LolB. The structures of LolA (PDB:1IWL) and LolB (1IWM) are strikingly similar, despite their dissimilar sequences. Both structures form a hydrophobic cavity consisting of 11 anti-parallel β-strand and three α-helices. These hydrophobic cavities bind the acyl chains of lipoproteins. The structures on the left were rotated by 90 degrees around the horizontal axis to give the structures on the right.

important for their functional differentiation. An extra loop comprising a short helix and a twelfth β-strand at the C-terminus of LolA prevents retrograde transfer of lipoproteins to the cytoplasmic membrane by inhibiting the interaction between LolA and phospholipids (Okuda *et al.*, 2008). The Arg residue at position 43 in the β2-strand of LolA forms hydrogen bonds with some residues in the lid, and therefore closes the cavity, whereas the cavity of LolB is

open. The LolA(R43L) mutant, in which Arg at position 43 is replaced by Leu, can accept lipoproteins from LolCDE, but cannot transfer them to LolB (Miyamoto *et al.*, 2001).
Therefore, an unusually large amount of the LolA–lipoprotein complex is accumulated in the periplasm (Taniguchi *et al.*, 2005). The hydrophobic interaction of LolA(R43L) with lipoproteins is as strong as that of LolB (Taniguchi *et al.*, 2005). This is why lipoproteins are not transferred from LolA(R43L) to LolB. In contrast, the interaction between wild-type LolA and lipoproteins is the weakest, allowing efficient transfer of lipoproteins to LolB. It is also noteworthy that the hydrophobic cavity of LolA is formed from aromatic residues, whereas that of LolB consists mainly of Leu and Ile, whose hydrophobic side chains are more flexible than those of aromatic residues (Takeda *et al.*, 2003). These differences are critical for the efficient one-way transfer of lipoproteins in an energy-independent manner.
It remains unknown how the three acyl chains are localized in LolA and LolB of *E. coli*, because the hydrophobic cavities of LolA and LolB are only large enough to accommodate a single acyl chain, i.e. not more. The crystal structure of *P. aeruginosa* LolA was recently shown to be almost the same as that of *E. coli* (Remans *et al.*, 2010). Additional hydrophobic patches were found on the surface of *P. aeruginosa* LolA. These patches might be the binding sites for other acyl chains of lipoproteins. However, LolA(R43L) crystals exhibited two structures, open and closed conformations. The hydrophobic cavity of the open conformation is larger (~1700 Å3) (Y. Oguchi and H. Tokuda, unpublished) than that of the closed form, which is essentially identical to the free form of wild-type LolA (Oguchi *et al.*, 2008). However, the hydrophobic cavity of wild-type LolA was found to undergo opening and closing upon the binding and release of lipoproteins, respectively (Oguchi *et al.*, 2008; Watanabe *et al.*, 2008). LprG, a lipoprotein of Gram-positive diderm *M. tuberculosis*, has a very similar structure to that of LolA. The hydrophobic cavity of LprG is large enough to accommodate three acyl chains (Drage *et al.*, 2010). Since its size was reported to be ~1500 Å3, the open conformation of LolA seems to be able to accommodate all three acyl chains inside the hydrophobic cavity.
In addition to that of LprG, the structures of the N-terminal domains of RseB, LppX and VioE are similar to those of LolA and LolB. RseB is a member of the envelope stress response system leading to the induction of σE expression. RseB is proposed to bind an unfolded lipoprotein in the hydrophobic cavity (Kim *et al.*, 2007). LppX is a lipoprotein of *M. tuberculosis* required for the translocation of complex lipids, the phthiocerol dimycocerosates (DIM), to the outer layer of the cell. Its large hydrophobic cavity was thought to be sufficient to accommodate a single DIM molecule (Sulzenbacher *et al.*, 2006).

These proteins seem to have similar function to LolA and LolB although the transported substrates differ. VioE is speculated to play a key role in the biosynthesis of violacein, a purple pigment with antibacterial and cytotoxic properties (Hirano *et al.*, 2008; Ryan *et al.*, 2008).

A soluble LolB derivative, mLolB, lacking an N-terminal lipid anchor is functional and able to replace LolB, although a higher amount of mLolB is required for normal growth (Tsukahara *et al.*, 2009). It was then found that mLolB expressed in the periplasm incorporates lipoproteins into not only the outer but also the cytoplasmic membrane or liposomes, indicating that mLolB has a lipid-targeting function, and does not distinguish the cytoplasmic and outer membranes. Therefore, the N-terminal membrane anchor of LolB is important for prevention of the mislocalization of lipoproteins to the cytoplasmic membrane. The Leu residue at position 68 is expected to play an important role in phospholipid targeting (Takeda *et al.*, 2003). Among the three major phospholipids, non-bilayer phospholipid PE is important for the LolB-dependent incorporation of lipoproteins (Tsukahara *et al.*, 2009). However, it remains largely unknown how mLolB discharges a cargo on the lipid surface.

How Lol proteins interact with each other

In order to elucidate how lipoproteins are transferred through Lol proteins, a photo-cross-linking technique developed by Schultz and his collaborators was applied to the Lol system (Okuda & Tokuda, 2009). This technique enables the introduction of an unnatural photo-reactive amino acid, *p*-benzoyl-phenylalanine (pBPA), into an amber (TAG) codon *in vivo* (Chin *et al.*, 2002; Ryu & Schultz, 2006; Wang *et al.*, 2006). A number of LolA or LolB mutants having an amber codon at desired positions were constructed. *E. coli* cells expressing such a mutant were irradiated with UV light for *in vivo* photo-cross-linking. Analyses of the cross-linked products revealed that LolA and LolB interact with each other at the entrances of their hydrophobic cavities. Moreover, the inside of the LolA cavity was found to interact with the outside of the LolB cavity. Lipoproteins were found to interact exclusively with the insides of the LolA and LolB cavities. NMR analyses revealed essentially the same mode of interaction between LolA and LolB (Nakada *et al.*, 2009). Taking these results together, it was proposed that lipoproteins are transferred from the cavity of LolA to that of LolB in a mouth-to-mouth fashion. In this model, the hydrophobic cavities of LolA and LolB are very close to each other, and therefore the transfer of lipoproteins occur smoothly in the direction of higher affinity, namely from LolA to LolB.

In vivo photo-cross-linking revealed that LolA interacts with LolC, but not with LolE, at the entrance of its cavity (Okuda & Tokuda, 2009). When pBPA was incorporated into lipoproteins, cross-linked products were obtained with LolE, but not LolC (unpublished data). From these results, it is speculated that LolC functions as a scaffold for LolA whereas LolE recognizes outer membrane-specific lipoproteins. LolC and LolE have similar topologies i.e. four membrane-spanning regions with a large loop exposed to the periplasm (Narita *et al.*, 2002; Yasuda *et al.*, 2009). Both proteins are essential for the growth of *E. coli*. These results indicate that LolC and LolE are functionally different, despite their structural and sequence similarities (26% identical). The periplasmic loops of LolC and LolE exhibit sequence similarity to LolA and/or LolB to some extent (~18% identical). It seems possible that these regions also have hydrophobic cavities. The transport of lipoproteins might be performed from LolCDE to LolA to LolB through the entrances, or mouths, of their hydrophobic cavities. The gene for LolE is only conserved in γ-proteobacteria i.e. not in other subdivisions (Narita & Tokuda, 2011). In these bacteria, the ABC transporters responsible for lipoprotein release are likely to comprise a LolCD homodimer, in which LolC can bind lipoproteins and interact with LolA.

The LolA–LolC interaction increases upon the binding of lipoproteins to LolCDE (Okuda & Tokuda, 2009). Interaction between the LolA–lipoprotein complex and LolB is not inhibited by an excess amount of free LolA (Watanabe *et al.*, 2007). The affinity of LolD for ATP increases upon lipoprotein binding (Ito *et al.*, 2006). Taking all these results together, the Lol system seems to have been designed to achieve the efficient one-way transfer of lipoproteins from the cytoplasmic to the outer membrane.

Lipoproteins involved in outer membrane biogenesis

Four major outer membrane components, phospholipids, LPS, β-barrel proteins, and lipoproteins, are hydrophobic and must be transported from the cytoplasmic membrane to the outer membrane via the hydrophilic periplasm. Then, the transported components must be correctly localized to the outer membrane. Since there is no energy source, such as ATP, in the periplasm, most of the transport reactions must take place in energy-independent manners. In addition to the Lol system described above, the machineries responsible for the transport and assembly of LPS and β-barrel proteins have been identified (Bos *et al.*, 2004; Ruiz *et al.*, 2008; Sklar *et al.*, 2007; Sperandeo *et al.*, 2007; Voulhoux *et al.*, 2003; Wu *et al.*, 2005). Various lipoproteins were found to play important roles in these transport-assembly reactions.

On the other hand, little is known about how phospholipids are transported to the outer membrane although a mechanism underlying the transport of phospholipids from the outer to the cytoplasmic membrane was reported (Malinverni & Silhavy, 2009).

LPS is flipped from the cytoplasmic side to the periplasmic side of the inner membrane by an ABC transporter, MsbA, and then transported to the outer leaflet of the outer membrane by the Lpt (Lipopolysaccharide transport) system. This system is composed of the cytoplasmic membrane ABC transporter LptBFG complex, bitopic membrane protein LptC, which forms a complex with LptBFG, periplasmic protein LptA, and the complex comprising β-barrel protein LptD and lipoprotein LptE (Sperandeo *et al.*, 2009). All factors comprising the Lpt system and MsbA are essential for the growth of *E. coli*. Although LPS is essential for most Gram-negative bacteria, *Neisseria meningitidis* cells grow in the absence of LPS biogenesis, and therefore mutants lacking any one of above factors are viable. Moreover, the deletion of *lptE* does not inhibit the transport of LPS to the cell surface, indicating that the function of LptE is indirect in this Gram-negative bacterium (Bos & Tommassen, 2011). On the other hand, *E. coli* LptE was reported to bind to LPS strongly (Chng *et al.*, 2010). LptE has a chaperone-like function in the LptD–LptE complexes of both *E. coli* and *N. meningitidis*. The mechanism of the Lpt system might not be the same among Gram-negative diderms.

The Bam (β-barrel assembly machinery) complex is required for the outer membrane assembly of β-barrel proteins in *E. coli* (Hagan et al., 2011). In addition, the periplasmic chaperone DegP, SurA and Skp are also involved in the targeting and assembly of β-barrel proteins. The Bam complex consists of one β-barrel protein, BamA, and four outer membrane-specific lipoproteins, BamB/C/D/E. A similar complex has been identified in *N. meningitidis*, although an outer membrane protein, RmpM, constitutes the complex instead of BamB (Volokhina *et al.*, 2009). Many components of the Bam complex are conserved in various bacteria (Anwari *et al.*, 2010; Gatsos *et al.*, 2008). BamA, a central component of the complex, is conserved in all Gram-negative bacteria. Among the four lipoproteins, only BamD is essential for the growth of *E. coli*, suggesting that BamD plays an important and direct role in the β-barrel protein assembly, while the functions of the respective Bam proteins remain to be clarified. The reconstitution of the Bam complex in proteoliposomes will help to solve these issues (Hagan *et al.*, 2010; Hagan & Kahne, 2011). The BamA orthologue is essential for the assembly of β-barrel proteins in the outer membrane of mitochondria (Gentle *et al.*, 2004; Voulhoux & Tommassen, 2004). Although some accessory components of the complexes differ between bacteria and mitochondria, which have no lipoproteins, the system for β-barrel protein assembly seems to have been evolutionary conserved.

As already mentioned, the Lol system mediates the sorting of lipoproteins to the outer membrane of *E. coli*, in which more than 90 species of lipoproteins are expressed (Tokuda, 2009). Most lipoproteins are predicted to be localized on the periplasmic side of the outer membrane. Although many lipoproteins have unknown functions, three essential lipoproteins, LolB, LptE and BamD, are involved in outer membrane biogenesis. Therefore, chemicals that inhibit the Lol system are expected to block the outer membrane biogenesis. Indeed, LolA has been shown to be a promising target of drugs (Pathania *et al.*, 2009).

Structural importance of lipoproteins

Bacterial lipoproteins are involved in a wide variety of cellular functions, such as the biogenesis and maintenance of cell surface structures, the transport of substrates, and drug efflux (Bernadac *et al.*, 1998; Clavel *et al.*, 1998; Ehrmann *et al.*, 1998; Nikaido, 1998). As discussed in later sections, some lipoproteins cause an inflammatory response in human host cells, and their lipid moieties account for the bioactivity (Ramesh *et al.*, 2003; Scragg *et al.*, 2000). Some lipoproteins have been shown to be functional without a lipid anchor (Chng *et al.*, 2010; Tsukahara *et al.*, 2009), although the activity of LolB decreases with a lack of the lipid anchor because it prevents the mislocalization of lipoproteins. The cytoplasmic membrane proteins have hydrophobic α-helical stretches, which cause the retention of the proteins in the cytoplasmic membrane. Proteins spanning the outer membrane have amphipathic β-strands, which do not cause the retention of the proteins in the cytoplasmic membrane. Posttranslational modification of proteins with a lipid anchor might have been evolved to generate proteins that can reach and associate with the outer membrane. The β-barrel proteins frequently form pores in the outer membrane, while many enzymatic functions are most likely due to lipoproteins and peripheral proteins.

It is known that perturbation of outer membrane biogenesis causes stress to *E. coli* cells and induces stress response systems. A LolA derivative, I93C/F140C, has two Cys residues in place of Ile at position 93 and Phe at position 140. Expression of this mutant in the absence of a reducing agent is lethal to *E. coli* because the hydrophobic cavity is closed by a disulfide bond, which activates the Cpx two-component system (Tao *et al.*, 2010). The Cpx system monitors the biogenesis of cell surface structures and plays an important role in the detection of misassembly of β-barrel proteins (Gerken *et al.*, 2010; Raivio & Silhavy, 2001). Since the Cpx pathway is generally induced by aberrant disulfide formation in the periplasm, the expression of LolA(I93C/F140C) might directly activate the pathway. Alternatively,

inhibition of the localization of essential lipoproteins by dominant negative LolA(I93C/F140C) causes activation of the Cpx system. The overexpression of YafY, a cytoplasmic membrane-specific lipoprotein, whose function is unknown, induces DegP expression through activation of the Cpx pathway, as the outer membrane lipoprotein NlpE does (Miyadai *et al.*, 2004). Since delipidated NlpE does not induce the expression of DegP, overproduction of NlpE may cause its mislocalization to the cytoplasmic membrane, and then activation of the Cpx system. Rcs phosphorelay is one of the envelope stress response systems in *E. coli*. The outer membrane-specific lipoprotein RcsF, which transmits the signal from the cell surface to the cytoplasmic membrane histidine kinase RcsC, activates the Rcs phosphorelay system when overexpressed. The observation that the overexpression of LolA activates the Rcs pathway (Chen *et al.*, 2001) also suggests that the Rcs system might monitor correct lipoprotein sorting. In any event, involvement in the envelope stress response systems is one of the functions of lipoproteins.

LIPOPROTEINS IN GRAM-POSITIVE DIDERM BACTERIA

Bacteria that exhibit Gram-positive staining belong to one of two divisions, *Actinobacteria* and *Firmicutes*. Formerly, these divisions were differentiated according to their G+C contents, i.e. high and low G+C Gram-positive bacteria, respectively. *Staphylococcus aureus*, *Micrococcus luteus* and *Bacillus anthracis* are typical members of the low G+C order *Firmicutes*. The cell envelope of these bacteria is characterized by a cytoplasmic membrane with a thick peptidoglycan layer and secondary glycopolymers, such as wall teichoic acid and lipoteichoic acid. Owing to the presence of a single membrane, these bacteria are classified as monoderm bacteria (Fig. 1c).

Within the division *Actinobacteria* and therein the order *Actinomycetales*, the genera *Mycobacterium*, *Corynebacterium*, *Rhodococcus* and *Nocardia* are grouped into the suborder *Corynebacterineae*. These genera share a common peptidoglycan structure and, in addition, are characterized by the presence of an outer membrane-like impermeable layer. Owing to the presence of a second membrane, these bacteria are considered to be diderm bacteria (Fig. 1b). Thus, from a general point of view, the cell envelopes of *Corynebacterineae* and *Proteobacteria* (Gram-negative bacteria) are similar. However, their structures are analogous rather than homologous. Major differences exist with respect to the chemical composition and, as a result of this with respect to permeability. For example, the lipid content of the cell

envelope of mycobacteria may represent up to 40% of the cellular dry mass, compared to only 10% in Gram-negative bacteria (Goren, 1979). The outer membrane of the Gram-negative bacterium *Pseudomonas aeruginosa* is notably impermeable. However, the mycobacterial outer membrane exhibits 10- to 100-fold lower permeability than that of *P. aeruginosa* (Jarlier & Nikaido, 1990).

Mycobacterium tuberculosis – *cell envelope composition and drug therapy*

Within the subfamily *Corynebacterineae*, lipoprotein function and biosynthesis have mainly been investigated in *Mycobacterium tuberculosis* and the model organism *Mycobacterium smegmatis*, respectively. *M. tuberculosis*, the causative agent of human tuberculosis annually causes 1.7 million deaths and about one-third of the world's population is latently infected (http://www.who.int/mediacentre/factsheets/fs104/en/index.html). *Mycobacterium bovis* bacille Calmette–Guerin (BCG), an attenuated derivative of *Mycobacterium bovis* (the causative agent of bovine tuberculosis), is applied as a tuberculosis live vaccine in high incidence countries. It efficiently protects against disseminated tuberculosis in childhood but only shows limited protection against adult lung tuberculosis (Kaufmann *et al.*, 2010). *M. tuberculosis* is an intracellular pathogen that survives and multiplies in macrophages (Russell, 2011). The biology of *M. tuberculosis* is intimately linked to its complex cell envelope. The structure, function and synthesis of the mycobacterial cell envelope have been extensively discussed recently, and readers interested in further details are referred to two excellent books (Cole, 2005; Daffe, 2008). The standard course therapy for tuberculosis is a four-drug regimen comprising isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA), and ethambutol (EMB) administered for 2 months, followed by INH and RIF alone for an additional 4 months. A dramatic increase in drug-resistant tuberculosis has been observed in recent years. In the past decade, novel antituberculosis drug candidates have been developed intensively (Cole & Riccardi, 2011). However, the tuberculosis drug pipeline is insufficiently filled. Therapy for drug-resistant tuberculosis requires the substitution of first line drugs by second line drugs and extended medication. Second-line drugs include fluoroquinolones [ofloxacin (OFL) and moxifloxacin (MOX)], aminoglycosides [amikacin (AMK) and kanamycin (KAN)], peptide antibiotics [viomycin (VIO) and capreomycin (CAP)] and thioamides [e.g. ethionamide (ETH)], cycloserine (CYC) and *para*-aminosalicylic acid (PAS)]. Two of the first-line drugs (INH and EMB), several second-line drugs (ETH and CYC) as well as drugs under development (Makarov *et al.*, 2009) inhibit the cell envelope

biogenesis. INH and ETH are pro-drugs that are activated within bacterial cells by catalase/peroxidase KatG and monooxygenase EthA, respectively (Almeida Da Silva & Palomino, 2011). Both INH and ETH form an adduct with nicotine adenine dinucleotide (NAD) and thereby inhibit the enoyl-acyl carrier protein (ACP) reductase component of a dissociable fatty acid synthase (FAS-II), which is involved in the synthesis of mycolic acid precursors. PA824 and OPC67683, two nitroimidazoles currently in phase II clinical trials, also interfere with mycolic acid biosynthesis (Cole & Riccardi, 2011). EMB interferes with arabinan biosynthesis. Most EMB-resistance associated mutations are found in the arabinosyltransferase EmbB. Benzothiazinone BTZ043 and dinitrobenzamide DNB are two promising antibiotic compounds under pre-clinical development. Both compounds inhibit the decaprenylphosphoryl-β-d-ribose 2′ epimerase, encoded by the *dprE1* and *dprE2* genes involved in D-arabinose synthesis. In particular, the enzymes transform decaprenylphosphoryl-d-ribose into decaprenylphosphoryl-d-arabinose, a precursor for the synthesis of mycobacterial cell-envelope polysaccharides arabinogalactan and lipoarabinomannan (LAM) (Cole & Riccardi, 2011). The number of antibiotics targeting cell envelope biogenesis emphasizes the importance of cell envelope biogenesis for *M. tuberculosis* viability and virulence.

The cell envelope structure and biogenesis of mycobacteria have been intensively studied by means of chemical analyses, microscopic techniques and genetic approaches during the past few decades. The structures surrounding the cytoplasm may be subdivided into several compartments, the cytoplasmic membrane, outer membrane-like structure (consisting of the cell wall core and extractable lipids) and the capsule.

The cytoplasmic membrane and peptidoglycan layer

The cytoplasmic membrane is mainly formed by phospholipids and proteins. The polar lipids are composed of hydrophilic head groups and fatty acid chains that usually consist of fatty acid residues of less than 20 carbon atoms. Palmitic acid (C16:0), octadecenoic acid (C18:1), and 10-methyloctadecanoic acid (tuberculostearic acid) are the major fatty acid constituents. The main phospholipids of the cytoplasmic membrane are phosphatidylinositol mannosides (PIM), PG, PE and CL (Daffe, 2008). The cytoplasmic membrane contains peripheral and integral membrane proteins as well as lipoproteins. It has a thickness of about seven nanometres, and it is supposed that it is structurally and functionally very similar to other bacterial cytoplasmic membranes (Daffe, 2008).

The cell wall core also called mycolyl-arabinogalactan-peptidoglycan (mAGP) complex is composed of peptidoglycan and molecules covalently linked to it such as arabinan and mycolylesters. This sacculus is an essential structure for mycobacteria and provides the bacteria with a formidable protective barrier against xenobiotics.

The mycobacterial peptidoglycan has some features that distinguish it from *E. coli* peptidoglycan. Its glycan chains are composed of alternating units of (β1→4)-linked *N*-acetylglucosamine (GlcNAc) and *N*-glycolylmuramic acid (MurNGlyc), whereas most other bacteria have *N*-acetylmuramic acid. A tetrapeptide (l-alanyl-d-isoglutaminyl-meso-diaminopimelyl-d-alanine) side chain substitutes for the carboxylic acid function of each muramic acid residue. The peptide chains are heavily cross-linked (70–80%) as compared to in *E. coli* (degree of cross-linking: 20–25%) (Mahapatra, 2005). Particularly in non-replicating *M. tuberculosis*, the non-classical 3→3 linkages between two diaminopimelyl residues (rather than the classical 4→3 linkages between the terminal alanyl- and diaminopimelyl residues) predominate in the transpeptide network (Gupta *et al.*, 2010). C-6 of some muramic acid residues form phosphodiester bonds with C-1 of α-D-GlcNAc, which in turn is (1→3) linked to an α-l-rhamno-pyranose (Rha*p*) residue. Rha*p* provides the 'linker unit' between peptidoglycan and the galactan of arabinogalactan (Mahapatra, 2005). The galactan consists of a linear chain of about 30 units of alternating 1→5 and 1→6 linked β-d-galactofuranose (gal*f*) residues. Two to three of the 1→6 linked galactan residues of the galactan chain are glycosylated at their C5 with an α-d-arabinofuranose (ara*f*) chain via an α 1→5 linkage. A tree-like arabinose structure is formed through the introduction of 1→2 or 1→3 branch points to the linear 1→5 linked stem. Mycolic acids are esterified to two-thirds of the terminal ara*f* (Mahapatra, 2005). Peptidoglycan and arabinogalactan are hydrophilic, while mycolic acids are hydrophobic and form part of the mycobacterial outer membrane-like structure.

Outer membrane-like structure

Mycolic acids are long chain fatty acids, both α-branched and β-hydroxylated. In *M. tuberculosis*, they contain 70–90 carbon atoms with various modifications of the acyl chains. According to a recent model, which is a modification of that originally proposed by Minnikin (Minnikin, 1982), mycolic acids form the inner leaflet of an outer membrane-like structure, along with phospholipids (Zuber *et al.*, 2008). The acyl chain of mycolic acids is not straight but folds back to fit the thickness of the outer membrane-like structure of 7–8 nm,

as observed on cryo-electron microscopy of vitreous sections (CEMOVIS) (Hoffmann *et al.*, 2008; Zuber *et al.*, 2008).

Besides forming the major part of the cell wall core component, mycolic acids also occur in unbound forms as esters of trehalose or glycerol and therefore are extractable with organic solvents. A huge variety of other complex lipids, such as phenolic glycolipids (trehalose monomycolate and trehalose dimycolate), phthiocerols and sulfolipids, interact with the covalently attached mycolic acids to build the outer leaflet of the outer membrane-like structure. The *M. tuberculosis* envelope lipids also contain phosphatidyl-*myo*-inositol mannosides (PIM) and their multiglycosylated counterparts, lipomannans (LM) and mannosylated lipoarabinomannans (ManLAM). In other Mycobacteria, LAM is modified differently. While PIM are found in the cytoplasmic membrane and also in the outer most layer, the capsule, the locations of LM and LAM remain controversial (Gilleron, 2008). Recent fractionation experiments on *M. smegmatis* indicated that LAM and LM are mainly co-localized with mycolic acids (Dhiman *et al.*, 2011). More than 400 lipid structures representing 15 classes of lipids have been identified in whole lipid extracts of *M. tuberculosis* (Sartain et al., 2011). The extraordinary ability of mycobacteria to degrade and synthesize lipids is reflected by the presence of 250 genes in the *M. tuberculosis* genome devoted to lipid metabolism, compared with 50 in *E. coli* (Cole et al., 1998).

The presence of a highly impermeable outer layer explains, at least partially, the natural resistance of mycobacteria to many hydrophilic antibiotics. However, the existence of such a hydrophobic barrier also poses a serious problem with respect to nutrient uptake. Nutrient uptake by mycobacteria and Gram-negative bacteria is facilitated by water-filled pore proteins (porins) present in the outer membrane (Niederweis *et al.*, 2010). Other proteins of the outer membrane include proteins of the PE and PPE families, two highly polymorphic sets of proteins characterized by the presence of Pro-Glu (PE) and Pro-Pro-Glu (PPE) motifs near the N-terminus (Mukhopadhyay & Balaji, 2011), and lipoproteins (Mawuenyega *et al.*, 2005).

The capsule mainly contains polysaccharides composed of a glycogen-like saccharides and d-arabino-d-mannan heteropolysaccharides, little lipids and proteins also being found in culture filtrates (Mahapatra, 2005).

Lipoprotein biosynthesis in mycobacteria

The genome of *M. tuberculosis* is of medium size (approximately 4 million basepairs) and encodes roughly 4000 genes. Approximately 2.5% of the open reading frames encode

lipoproteins (Sutcliffe & Harrington, 2004). Lipoprotein biosynthesis in mycobacteria follows the general lipoprotein biosynthesis pathway, as described above for *E. coli* (Rezwan *et al.*, 2007b; Sutcliffe & Harrington, 2004). Most preprolipoproteins are secreted via the general secretory pathway (*sec*) in a *secA1*- or *secA2*-dependent manner (Gibbons *et al.*, 2007); however, a significant number of preprolipoproteins are predicted to be translocated across the cytoplasmic membrane via the twin arginine transport pathway (*tat*) (McDonough *et al.*, 2005). After export, a conserved motif, the lipobox motif [LVI]-[ASTVI]-[GAS]-C (Babu *et al.*, 2006), is recognized by the membrane-bound pre-prolipoprotein diacylglycerol transferase (Lgt; Rv1614). A thioether linkage between Cys of the lipobox and diacylglycerol is formed. Attachment of the diacylglycerol anchor triggers lipoprotein signal peptidase (LspA; Rv1539)-dependent cleavage of the immature lipoprotein. This reaction generates a free amino group in the Cys carrying the thioether-linked diacylglycerol. It has long been assumed that mycobacterial lipoproteins are *N*-acylated. However, direct evidence of *N*-acylation of mycobacterial lipoproteins was only provided recently (Tschumi *et al.*, 2009).

Mycobacterial Lnt

A homologue of *E. coli* apo-lipoprotein *N*-acyltransferase (Lnt) is present in the genome of *M. tuberculosis* and in a wide variety of other GC-rich Gram-positive bacteria, but is absent from the genomes of low GC Gram-positive monoderms (Fig. 5). In *M. tuberculosis*, open reading frame Rv2051c is annotated as *ppm1*, for a polyprenol monophosphomannose (Ppm) synthase, which transfers mannose from GDP-mannose to endogenous polyprenol phosphate. Polyprenol phosphate is a metabolic intermediate in the synthesis of mycobacterial cell envelope constituents LM and LAM. The Ppm synthase domain is only encoded by the 3′ part of open reading frame 2051c. The 5′ part of this open reading frame encodes an amino terminal domain that exhibits considerable similarity to that of *E. coli* Lnt. *M. smegmatis*, a fast growing non-pathogenic *Mycobacterium*, which is tractable to genetic manipulation, is often used as a model organism for elucidating mycobacterial metabolic pathways. Heterologous expression of *M. tuberculosis* lipoprotein LppX in *M. smegmatis* and subsequent mass spectrometry analyses provided direct proof of *N*-acylation of mycobacterial lipoproteins (Tschumi *et al.*, 2009).

In *M. smegmatis*, orthologues of the two domains of *M. tuberculosis* Rv2051c are encoded by two distinct open reading frames, Msppm1 and Msppm2. Of these, Msppm2 (Msmeg3863) corresponds to Lnt. Targeted gene inactivation of Msppm2 results in a mutant that is unable to modify lipoprotein LppX with an *N*-acyl residue. Complementation of the mutant strain with

either *M. smegmatis lnt* (Msmeg3863) or *M. tuberculosis* Rv2051c restores the capability of *N*-acylation of LppX (Tschumi *et al.*, 2009). Subsequently, expression and analyses of recombinant *M. tuberculosis* lipoprotein LprF in wild-type *M. smegmatis* and its *Δlnt* strains indicated that *lnt*-dependent *N*-acylation is not restricted to LppX but represents a general modification pathway for mycobacterial lipoproteins (Brülle *et al.*, 2010).

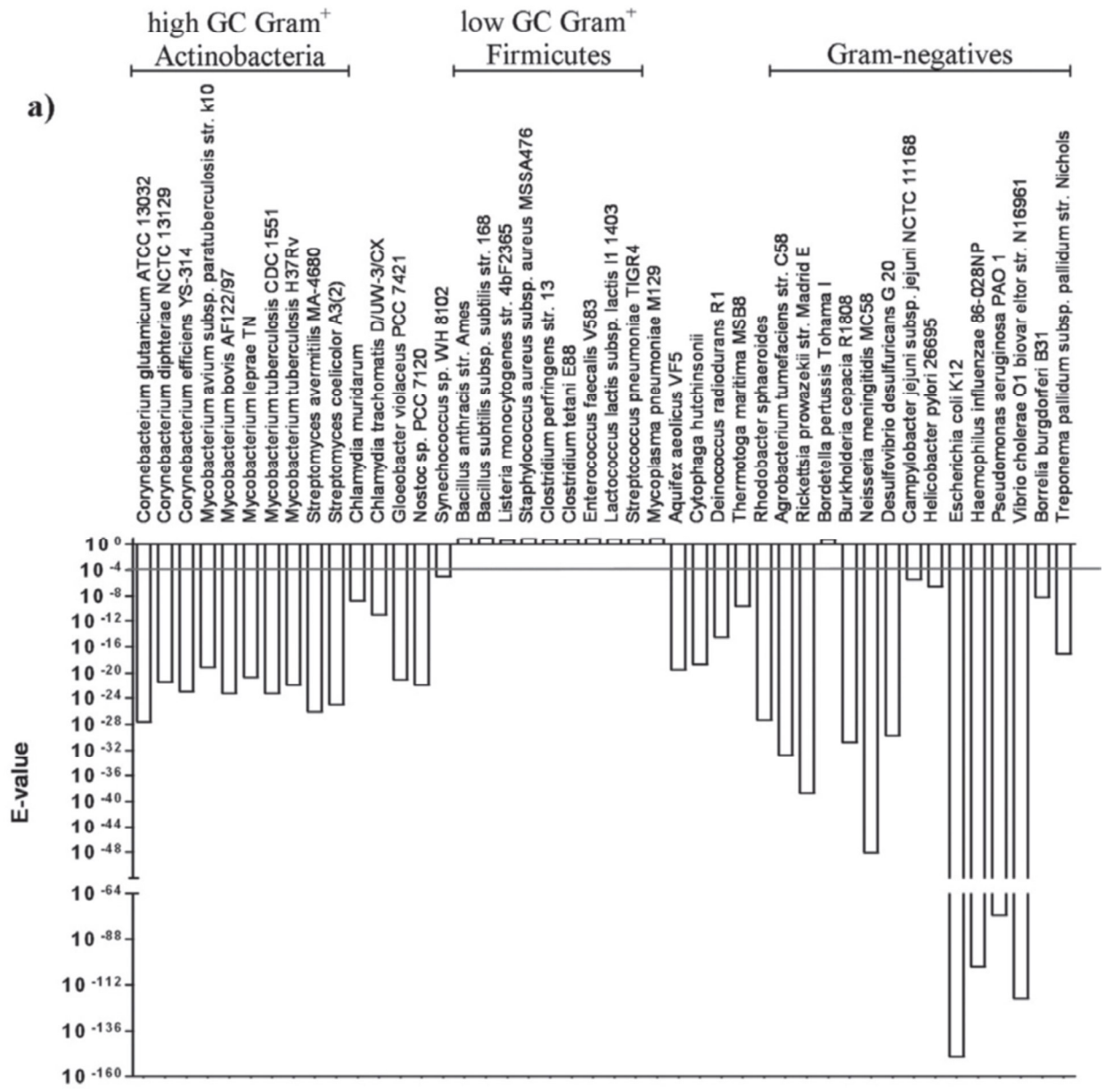


Fig. 5 Distribution of Lnt in bacteria. *E. coli* Lnt was used as a query to identify homologues on the National Centre of Biotechnology Information BLASTp server (http://www.ncbi.nlm.nih. gov/sutils/genom_table.cgi). The sequence filtering option was switched off, and the expected value was set at 10 and the cut-off value at 10–4. (Figure adapted from Tschumi *et al.*, J. Biol. Chem., 2009.)

The structures of the membrane anchor of the two recombinant lipoproteins, LppX and LprF, have been resolved at the molecular level. In both cases, the thioether-linked diacylglycerol carries one esterified palmitic acid and one esterified tuberculostearic acid. Whether the tuberculostearic acid is at the *sn*-1 or *sn*-2 position has not been determined yet. The *N*-acyl residue is mainly derived from palmitic acid. Thus, the distribution of fatty acids in mycobacterial lipoproteins reflects the composition of fatty acids in membrane phospholipids. In the genome of *M. bovis*, a second open reading frame, MB2285c, encoding a protein of 502 amino acids exhibits considerable homology to that of *E. coli* Lnt. Two of the amino acids of the catalytic triad (E267 and K335; *E. coli* numbering) are conserved, while the third residue (C387) is Ser (Vidal-Ingigliardi *et al.*, 2007). Most interestingly, the corresponding open reading frame in *M. tuberculosis* is split into two and the putative catalytic triad lies on two adjacent open reading frames. Of these, Rv2262c encodes a protein of 360 amino acids and Rv2261c encodes a protein of 140 amino acids. It remains to be elucidated whether or not proteins MB2285c and Rv2262c in combination with Rv2261c function as apolipoprotein *N*-acyl transferases, and whether or not they have the same specificity as Ppm1 (Rv2051c).

Lipoprotein biosynthesis and virulence

Lipoproteins have been shown to play a direct role in the interaction of bacterial pathogens with their hosts (Kovacs-Simon *et al.*, 2011). The virulence-related functions of lipoproteins include colonization, invasion, evasion of host defence, and immunomodulation. Additionally, lipoproteins have essential functions in general bacterial physiology, e.g. in nutrient acquisition or cell envelope biogenesis, and thereby indirectly affect host–pathogen interactions. The lipoprotein biosynthesis pathway is an important virulence factor of *M. tuberculosis*.

An *M. tuberculosis lspA* knock-out mutant exhibited a 3–4 log reduced number of colony forming units in an animal model of tuberculosis. Mice infected with the *lspA* mutant strain hardly show any lung pathology (Fig. 6) and the bacteria do not spread to secondary organs (Rampini *et al.*, 2008; Sander *et al.*, 2004). *M. tuberculosis* mutants deficient in *lgt*, the gating enzyme of the lipoprotein biosynthesis pathway, have not been described yet and high density transposon site hybridization mutagenesis (TRASH) and targeted mutagenesis experiments support the view that this gene is essential in *M. tuberculosis* (Sassetti & Rubin, 2003; Tschumi et al., 2012). Although *lspA* had also been predicted to be essential, such a mutant could be generated (Banaiee *et al.*, 2006; Sander *et al.*, 2004). An *M. tuberculosis ppm1/lnt*

mutant was isolated on screening for mutants unable to grow in an acidified medium. Detailed characterization of the mutant remains to be reported (Vandal *et al.*, 2008).

Mycobacteria and Gram-negative bacteria are diderms. Despite their similarities, these bacteria differ with respect to the essentiality of lipoprotein synthesis genes; all genes involved in lipoprotein synthesis and transport are essential for *E. coli* (Okuda & Tokuda, 2011).

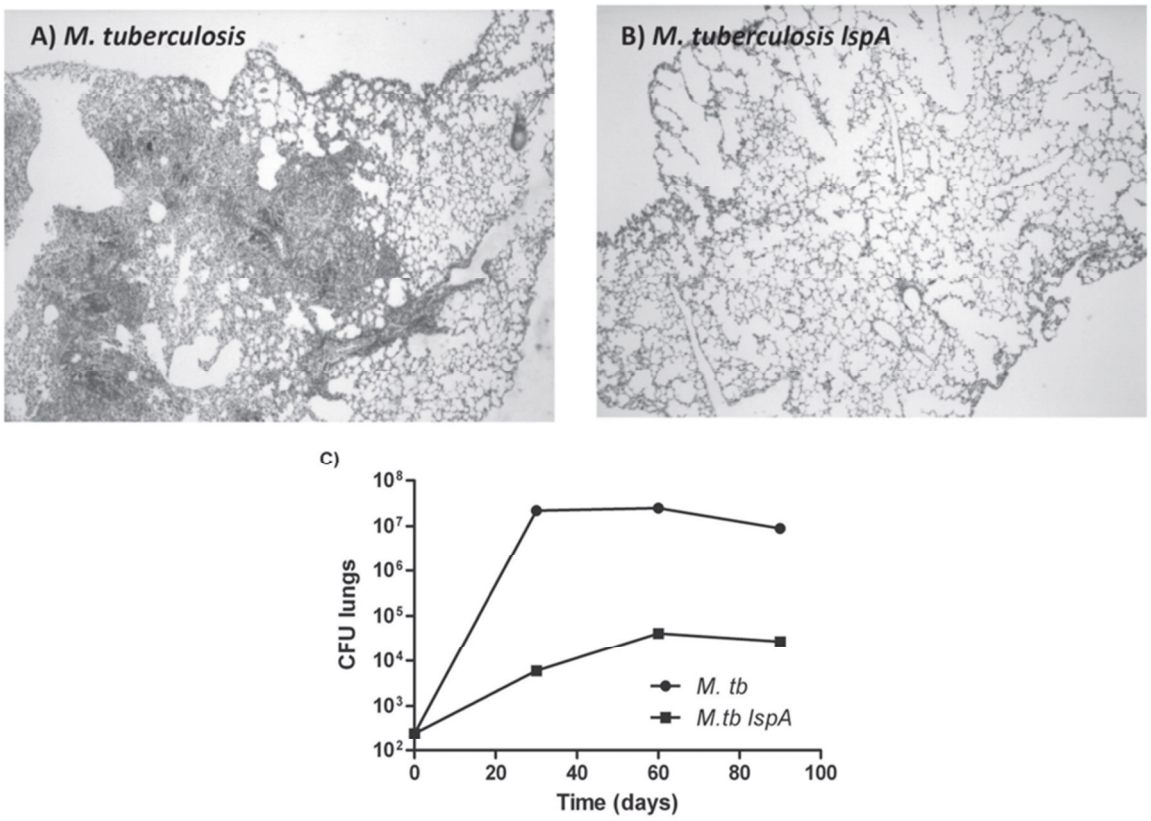


Fig. 6 Virulence attenuation of *M. tuberculosis* lipoprotein biosynthesis mutant *lspA*. The lung pathology of BALB/c mice infected with *M. tuberculosis* (A) and *M. tuberculosis lspA* mutant (B) at 42 days after aerosol infection is shown, representative haematoxylin/eosin-stained views of lung cross sections (total magnification x 40) being shown. Note the almost complete absence of inflammatory infiltration in the lungs infected with the mutant strain lacking functional LspA. (C) Growth of *M. tuberculosis* and *M. tuberculosis lspA* in the lungs of CBA/J mice over a 90-day period after intranasal infection (figures from Sander *et al.*, 2004).

One reason for the viability of mycobacterial mutants defective in lipoprotein biosynthesis (and other monoderm Gram-positive bacteria) could be that some lipoprotein precursors retain functionality (Hutchings *et al.*, 2009). Globomycin, a cyclic peptide antibiotic produced by *Streptomyces*, inhibits LspA-dependent processing of lipoproteins in both Gram-negative and Gram-positive bacteria (Inukai *et al.*, 1978). A bactericidal effect of globomycin on Gram-

negative bacteria but not on Gram-positive bacteria is consistent with the essentiality of LspA in the former bacteria, and a non-essential role in the latter ones. Globomycin exhibits a growth inhibitory effect on *M. tuberculosis*; however, this effect is not due to LspA inhibition since processing of lipoprotein Mpt83 is not affected, and an *M. tuberculosis lspA* mutant is as sensitive to globomycin as the parental strain (Banaiee *et al.*, 2007). Antibacterial activity of globomycin derivatives in monoderm Gram-positive bacteria, in which *lspA* is not essential, supports the view that globomycin may have additional targets (Kiho *et al.*, 2003; Kiho *et al.*, 2004).

Glycosylation of mycobacterial lipoproteins

Mycobacterial lipoproteins often contain additional post-translational modifications, besides the membrane anchor. Several *M. tuberculosis* lipoproteins (LpqH, SodC, LppX and LprF) have been shown to be modified with one or more glycosyl residues at their N-termini (Brülle *et al.*, 2010; Herrmann *et al.*, 1996; Sartain & Belisle, 2009; Tschumi *et al.*, 2009). *O*-glycosylation occurs at Thr and Ser residues, respectively (Sartain & Belisle, 2009). Glycosylation increases protease resistance, while the exact molecular nature of the glycosylation, as well as its function, remain largely undefined.

Transport and localization of lipoproteins in mycobacteria

Functional and structural investigations have suggested the localization of at least some *M. tuberculosis* lipoproteins in the mycobacterial outer membrane: *M. tuberculosis* lipoproteins have been described as adhesins, immune electron microscopy techniques have demonstrated their surface localization, and fractionation experiments have indicated co-localization with the fraction containing mycolic acids. Analyses of cytoplasmic membrane and outer membrane fractions allowed the identification of 21 lipoproteins (Mawuenyega *et al.*, 2005). More recently, 24 lipoproteins were identified in the detergent fraction of cell extracts (Wolfe *et al.*, 2010). The outer membrane transport of lipoproteins in Gram-negative bacteria depends on the localization of the lipoprotein Lol transport system and it has been suggested that sorting pathways for lipoproteins are critically important in all diderm bacteria (Okuda & Tokuda, 2011). The presence of lipoproteins in the outer membrane of mycobacteria suggests the existence of a lipoprotein transport system. However, its components remain to be identified. Mycobacteria release membrane vesicles *in vitro* and

during infection. Interestingly, the membrane vesicles of pathogenic mycobacteria are enriched in lipoproteins and contribute to virulence (Prados-Rosales *et al.*, 2011).

Mycobacterial lipoproteins in health and disease

Manipulation of the immune system

Lipoproteins affect both innate and adaptive immunity, and have been identified as major antigens of *M. tuberculosis*. Mycobacterial lipoproteins trigger the activation of humoral and cellular immune responses to mycobacteria. Some lipoproteins (LpqH, PstS1) induce a protective immune response, while others are deleterious as to protection (LprG) (Hovav *et al.*, 2003). Successful immune evasion of *M. tuberculosis* has partly been attributed to Toll-like receptor 2 (TLR2)-dependent inhibition of antigen processing and presentation (Drage *et al.*, 2009; Harding & Boom, 2010). Lipoproteins (from different bacteria including mycobacteria) are potent agonists of TLR2. Although TLR signalling enhances both innate and adaptive immune responses, it can also down regulate some immune functions. TLR2, in particular, has been implicated in the down regulation or deviation of the immune response through the induction of interleukin 10 and T helper 2 cell or regulatory T cell responses. Prolonged TLR signalling might constitute homeostatic feedback regulation that limits the extent of the induced responses. TLR2 agonist activity has been demonstrated for several *M. tuberculosis* lipoproteins including LpqH, LprA, LprG and PstS1, and also for mycobacterial glycolipids such as PIM, LM and LAM. Recognition of these molecules may depend on different TLR2-coreceptors and accessory receptors (Drage *et al.*, 2009). *De novo* MHC class II antigen processing and presentation are inhibited by prolonged signalling with agonists of TLR2. Down-regulation of antigen presentation is not specific to *M. tuberculosis*, but it could be especially pronounced during infection with *M. tuberculosis*. *M. tuberculosis* survives and multiplies in an early endosomal compartment by arresting phagosome maturation. Thus, *M. tuberculosis* is persistently co-localized with TLRs. The prolonged residence of *M. tuberculosis* in phagosomes and the abundance of cell envelope ligands for TLR2 that are released from viable mycobacteria, and that can transfer out of the phagosomes and out of the infected cells into the neighbour cells, provide ample opportunity for TLR2 signalling in the phagosomes or on the surface of antigen-presenting cells. Thus, TLR2-dependent suppression of MHC class II expression and processing could be particularly relevant in *M. tuberculosis* (Harding & Boom, 2010).

Mycobacterial lipoproteins

The genome of *M. tuberculosis* encodes approximately 100 lipoproteins. Some of these proteins have annotated functions, which have been reviewed previously (Rezwan *et al.*, 2007a; Sutcliffe & Harrington, 2004). Despite some progress, many of the open reading frames corresponding to lipoproteins still do not have annotated functions. Transposon site hybridization mutagenesis of *in vitro* grown *M. tuberculosis* mutants has suggested that several lipoprotein genes, i.e. *lpqW*, *lprB*, *lpqF* and *dppA*, are essential (Sassetti *et al.*, 2003). Additional lipoprotein genes, i.e. *lpqY*, *lpqZ*, *lprK*, *lpqT*, *lprG*, *lppX* and *lprN*, were shown to be required for virulence on corresponding *in vivo* screening (Sassetti & Rubin, 2003). Besides these genome-wide screenings, a variety of individually generated *M. tuberculosis* mutants deficient in certain lipoprotein genes were attenuated and several lipoproteins were characterized in more detail. SodC, LpqH, PstS2, PstS3, LppX, LprG, LpqN, LprC, LppK and LpqG were found to be abundantly expressed (Malen *et al.*, 2010). Among a large number of *M. tuberculosis* lipoproteins, some lipoproteins including ones most recently found are discussed here. The *M. tuberculosis* database TubercuList (http://tuberculist.epfl.ch) hosted by the Institute of Global Health, École Polytechnique Fédérale de Lausanne, Switzerland, provides an excellent update for each *M. tuberculosis* open reading frame and is linked to other mycobacterial databases.

LprG (Rv1411c)

LprG has been identified in mycobacterial cell envelope fractions by means of various techniques. *lprG* is a non-essential gene of *M. tuberculosis*, but is required for the survival of *M. tuberculosis* in murine macrophages and for full virulence in mice. The gene is co-transcribed with Rv1410c, a protein of the major facilitator superfamily of small molecules. Inactivation of the homologous gene in *M. smegmatis* revealed a role of each of these genes in ethidium bromide resistance. Moreover, both mutants were deficient in sliding motility and exhibited an altered colony morphology. These observations suggest a role of the *lprG*-operon in the cell envelope function (Farrow & Rubin, 2008). *M. tuberculosis* LprG as well as other lipoproteins are potent TLR agonists. Surprisingly, a non-acylated LprG still exhibited TLR2 stimulation, despite the absence of a membrane anchor. The remaining TLR2-stimulating activity of LprG was attributed to glycolipid binding. The crystal structure of LprG revealed a binding pocket that could accommodate lipids with three acyl chains. Based on structural data, a role of LprG in outer membrane transport of mycobacterial lipoproteins has been considered. LAM, LM and PIM were co-purified with recombinant LprG. In contrast,

mycobacterial lipoproteins can not be co-purified with LprG. These observations suggest a primary role of LprG in lipoglycan transport rather than in lipoprotein transport (Drage *et al.*, 2010). As mentioned in 'Lipoproteins in Gram-negative diderm bacteria' above, the crystal structures of LprG and LppX are very similar to those of LolA/LolB.

LppX (Rv2945c)

M. tuberculosis LppX, similar to LprG, is also involved in the transport of complex lipids. This protein has a hydrophobic cavity that binds phthiocerol dimycocerosates (DIM) and delivers them to the outer membrane (Sulzenbacher *et al.*, 2006). LppX orthologues are restricted to mycobacteria capable of synthesizing DIM, suggesting that their transport function is selective. An *M. tuberculosis* mutant deficient in *lppX* is attenuated in a mouse model of tuberculosis (Camacho *et al.*, 1999).

Rpf B (Rv1009)

The genome of *M. tuberculosis* encodes five resuscitation-promoting factor (Rpf) proteins, one of them, RpfB, being a lipoprotein. The Rpf proteins have a c-type lysozyme-like fold, and are predicted to cleave the glycosidic bond between *N*-acetyl glucosamine and *N*-acetyl muramic acid in peptidoglycan. Characterization of single and multiple *rpf* mutants produced conflicting results with respect to the importance of individual genes in resuscitation, which was due to the use of different resuscitation and virulence models, respectively (Chao & Rubin, 2010).

LppA (Rv2543)

LppA is a lipoprotein confined to mycobacterial pathogens. One copy of its gene is present in the genome of *M. tuberculosis* H37Rv, while two copies (LppA and LprR) are present in the genome of *M. tuberculosis* CDC1551. LppA is highly homologous to its downstream gene, LppB (Rv2544), with which it shares 90% amino acid identity. The structure of LppA has been solved at 2 Å resolution. However, the lack of sequence and structural homologues of LppA hinders any functional assignments (Grana *et al.*, 2010).

LpqB (Rv3224)

LpqB (Rv3244c) is one of the 233 conserved signature proteins of the *Actinobacteria* (Gao et al., 2006). The *lpqB* gene is located downstream of the essential signal transduction system, MtrAB. LpqB is part of a three-component system that co-ordinates cytokinetic and cell

envelope homeostatic processes in mycobacteria, and directly interacts with the extracellular domain of MtrB. An *M. smegmatis lpqB* mutant was isolated during screening for multiple antibiotic resistance-defective phenotypes. The mutant exhibited increased cell–cell aggregation, and severe defects in surface motility and biofilm growth (Nguyen *et al.*, 2010).

LpqM (Rv0419)

Screening of *M. smegmatis* mutants deficient in conjugal DNA transfer revealed *lpqM*. LpqM is a putative lipoprotein with a metalloprotease signature sequence (HExxH). Both *M. smegmatis* and *M. tuberculosis lpqM* (Rv0419) restore DNA transfer. Secretion and lipidation of LpqM are mandatory for complementation of the phenotype. Expression of functional LpqM in the donor strain is sufficient for proficient transfer (Nguyen *et al.*, 2009). So far, a substrate of LpqM has not been identified.

LpqY(Rv1235)-SugA-SugB-SugC operon

Five putative carbon uptake permeases are encoded in the genome of *M. tuberculosis*. Of these, LpqY-SugA-SugB-SugC encodes an ABC transporter highly conserved in mycobacteria. It comprises periplasmic sugar-binding lipoprotein LpqY, ATP-binding protein SugC, and transmembrane proteins SugA and SugB. This ATP-binding cassette transporter is highly specific for the uptake of disaccharide trehalose. Disaccharide trehalose is not present in mammals. However, trehalose is released as a byproduct of the biosynthesis of the mycolic acids by the mycolyltransferase antigen 85 complex. The antigen 85 complex transfers the lipid moiety of the glycolipid trehalose monomycolate (TMM) to arabinogalactan or another molecule of TMM, yielding trehalose dimycolate. These reactions also lead to the concomitant extracellular release of the trehalose moiety of TMM. Disruption of retrograde trehalose transport and thus impairment of trehalose recycling attenuates *M. tuberculosis* (Kalscheuer et al., 2010).

SodC (Rv0432)

SodC encodes a Cu,Zn-dependent superoxide dismutase and is annotated as a putative lipoprotein based on the presence of a lipoprotein lipid attachment site and radioactive labelling upon heterologous expression in *E. coli*. It detoxifies reactive oxygen intermediates and thereby contributes to the survival of *M. tuberculosis*, particularly in activated macrophages. The crystal structure of SodC has been solved by X-ray crystallography. The enzyme is glycosylated at its N-terminus, and is localized to the cytoplasmic and outer

membrane fractions. However, its lipidation has recently been questioned due to the results of fractionation and mutagenesis experiments combined with 2D gel electrophoresis and mass spectrometry (Sartain & Belisle, 2009).

LpqH (Rv3763)

The 19-kDa antigen (LpqH, Rv3763) of *M. tuberculosis* has been recognized as an immunodominant lipoprotein and often has been used as a mycobacterial model lipoprotein, particularly as a TLR2 agonist. LpqH is glycosylated and functions as an adhesin by binding to the mannose receptor of monocytic cells (Diaz-Silvestre *et al.*, 2005). An *M. tuberculosis lpqH* knock-out mutant is attenuated in IFN-γ activated monocyte-derived macrophages and in mice. The *lpqH* mutant nearly does not multiply in C57BL/6 mice or even in IFN-γ deficient mice. When applied as a live vaccine, the *M. tuberculosis lpqH* mutant exhibits similar protective efficacy to BCG against an aerosol challenge with *M. tuberculosis* (Henao-Tamayo et al., 2007).

Lipoprotein biosynthesis in Streptomyces

Streptomyces also has a GC-rich genome, exhibits Gram-positive staining and is a member of the order *Actinomycetales*. *Streptomyces* species are characterized by a complex secondary metabolism. They produce a wide variety of antibiotics, e.g. streptomycin and kanamycin, and antifungal and anti-parasitic drugs, and also immunosuppressants. Some of the 550 described species are pathogens of humans or plants. Unlike their relatives from the subfamily *Corynebacterineae*, they lack an outer membrane-like structure. *Streptomyces* export lipoproteins via the *sec* and *tat* pathways (Shruthi *et al.*, 2010; Thompson *et al.*, 2010). Interestingly, *Streptomyces coelicolor* encodes two functional Lgts (*lgt1* and *lgt2*). Both *lgt* genes could be disrupted separately, but a double deletion mutant was not viable. Likewise an *lspA* knock-out mutant of this species could only survive upon acquisition of suppressor mutation (Thompson *et al.*, 2010). In contrast, *lgt* and *lspA* deletion mutants of the plant pathogen *Streptomyces scabies* are viable. These mutants are defective in vegetative growth and spore development but only moderately attenuated in virulence (Widdick *et al.*, 2011). *lgt*-deficient mutants release some (non-lipidated) lipoproteins into the culture supernatant through shedding or shaving, while other lipoproteins are either not expressed or are degraded (Widdick *et al.*, 2011). The model lipoprotein SCO3484 was shown to be modified with three fatty acids: two ester-bound and one amide-bound fatty acid are attached to the N-terminal

Cys. The length of the fatty acids varies between 15 and 18 carbon atoms. All *Streptomyces* species encode two Lnt-homologues. The *N*-acylation of SCO3484 is completely abolished in an *S. scabies lnt1* deletion mutant and diminished in an *lnt2* mutant. These data indicate that Lnt1 is a functional lipoprotein *N*-acyltransferase, while the role of Lnt2 remains elusive (Widdick *et al.*, 2011).

LIPOPROTEINS IN GRAM-POSITIVE MONODERM BACTERIA

Low GC-content Gram-positive bacterial species such as *Bacillus*, *Clostridium*, *Enterococcus*, *Listeria*, *Staphylococcus*, *Streptococcus* and cell wall-less mycoplasma have a single cellular membrane and are classified as monoderm bacteria. An *E. coli* Lnt homologue was not found in their genomes and therefore they had been expected to have only diacylated lipoproteins (Hutchings *et al.*, 2009; Kovacs-Simon *et al.*, 2011; Sutcliffe & Russell, 1995). However, earlier biochemical studies suggested that the lipoproteins of certain *Bacillus* and *Staphylococcus* species could be *N*-acylated (Navarre *et al.*, 1996; Nielsen & Lampen, 1983). It is important to determine whether or not the low-GC-content monoderm bacteria have an Lnt-like enzyme, and to clarify the exact structures subjected to lipid modification in these bacteria (Kovacs-Simon *et al.*, 2011; Schmaler *et al.*, 2010). To address these issues, Lee and his collaborators recently provided biochemical evidence that staphylococcal lipoproteins are *N*-acylated triacyl forms (Asanuma *et al.*, 2011; Kurokawa *et al.*, 2009). Serebryakova *et al.* (2011) also reported that lipoproteins of *Acholeplasma laidlawii*, a mycoplasma strain, are triacylated. These studies indicate that staphylococcal and some mycoplasma species must have another type of Lnt whose structure is distinct from that in *E. coli*. Furthermore, novel lipopeptide structures have been identified in low GC content Gram-positive bacteria (Kurokawa *et al.*, 2012b). Their detailed structures, putative biosynthetic pathways, and biological functions are discussed in this section.

N-terminal structures of Staphylococcus aureus lipoproteins

Hantke and Braun (1973) reported the structure of the major outer membrane lipoprotein, Lpp, of *E. coli* in 1973. In this classical study, the Lpp structure was clarified by a combination of several methods including chemical degradation, radioisotope incorporation and chemical synthesis. However, detailed structural analyses by means of these methods are

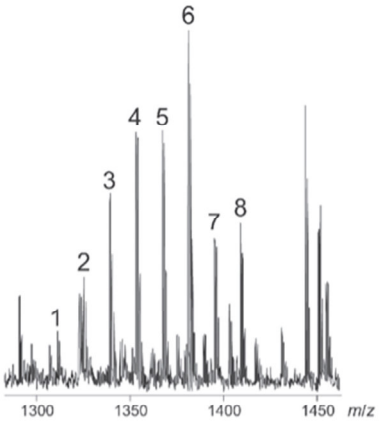
not always easy because large amounts of purified lipoproteins are required. In contrast, mass spectrometric (MS) analysis, which was developed recently (Murphy & Gaskell, 2011), enables the direct analysis of the acylated states of bacterial lipoproteins. Furthermore, this method quantitatively reveals the lipid compositions of bacterial lipoproteins. Indeed, MS-based analysis has been used for the determination of lipoprotein/lipopeptide structures and their lipid compositions (Hashimoto *et al.*, 2004; Muhlradt *et al.*, 1997).

The MS-based strategy was applied to determine the molecular lipid moieties attached to the N-terminal Cys residue of *S. aureus* lipoproteins (Kurokawa *et al.*, 2009). *S. aureus* is a major human pathogen causing evasion of the host's protective immune responses and is known to affect human immunity (Foster, 2005). Furthermore, antibiotic-resistant strains such as methicillin- and vancomycin-resistant *S. aureus* ones are emerging worldwide (Lowy, 1998). Importantly, *S. aureus* lipoproteins/lipopeptides are recognized by a host pattern recognition receptor, Toll-like receptor 2 (TLR2), and activate the TLR signalling pathway, leading to the production of pro-inflammatory cytokines and the establishment of adaptive immunity (Iwasaki & Medzhitov, 2010; Takeuchi & Akira, 2010). Therefore, it is important to determine the exact structures of lipoproteins as immunostimulatory ligands in order to further understand host–pathogen interactions.

A lipoprotein-enriched fraction was prepared from *S. aureus* RN4220 cell lysates by means of Triton X-114 phase partitioning and then separated by SDS-PAGE (Kurokawa *et al.*, 2009). A 33-kDa protein band was identified as lipoprotein SitC on LC (liquid chromatography)-MS/MS analysis. The N-terminal peptides of SitC were prepared by in-gel digestion with trypsin and then analysed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). Surprisingly, a series of eight mass peaks exhibiting 14 mass differences corresponding to the theoretical *m*/*z* of *N*-acyl-*S*-(diacylglyceryl)-cysteinyl peptides were detected (peaks 1–8, Fig. 7a). These peaks consisted of those of three fatty acids and increasing numbers of methylene (-CH2-) groups. Consistent with this, Edman degradation failed to reveal the N-terminal amino acid sequence. These results indicate that *S. aureus* SitC is a triacylated lipoprotein (Kurokawa *et al.*, 2009). In contrast, Tawaratsumida *et al.* (2009) reported that another lipoprotein, SA2202, of *S. aureus* SA113 had a diacylated (dipalmitoylated) *N*-terminus, based on the results of MS/MS analyses. Thus, the *N*-acylation of *S. aureus* lipoproteins is controversial. Asanuma *et al.* (2011) then examined whether the α-amino group of the *S*-diacylglyceryl Cys residue of *S. aureus* lipoproteins is acylated or not. Highly purified lipopeptides generated from SitC were analysed by MALDI-TOF MS and revealed to be triacylated (Fig. 7b, left). In this study, the N-terminal SitC lipopeptides were

treated with lipoprotein lipase (LPL) and then analysed. LPL specifically hydrolyzes the ester-linked fatty acids of bacterial lipoproteins. This treatment generated a new series of 14-Da interval peaks corresponding to the diacylglyceryl CGTGGK SitC lipopeptide (Fig. 7b, lower right). After long-term treatment with LPL, peaks corresponding to monoacylglyceryl CGTGGK SitC lipopeptides (Fig. 7b, upper right) were also generated because of the release of two *O*-esterified fatty acids. Taken together, these results indicate that SitC is modified by two *O*-esterified fatty acids and one LPL-resistant fatty acid.

a



Pos.	*m/z* obtained	Corresponding structure & its *m/z*	
1	1310.9798	Pam_3-CGTGGK,	1310.9604
2	1324.9886	(Pam_3+CH_2)-CGTGGK,	1324.9760
3	1338.9963	[Pam_3+(CH_2)$_2$]-CGTGGK,	1338.9917
4	1353.0136	[Pam_3+(CH_2)$_3$]-CGTGGK,	1353.0073
5	1367.0318	[Pam_3+(CH_2)$_4$]-CGTGGK,	1367.0230
6	1381.0498	[Pam_3+(CH_2)$_5$]-CGTGGK,	1381.0386
7	1395.0704	[Pam_3+(CH_2)$_6$]-CGTGGK,	1395.0543
8	1409.0792	[Pam_3+(CH_2)$_7$]-CGTGGK,	1409.0699

b

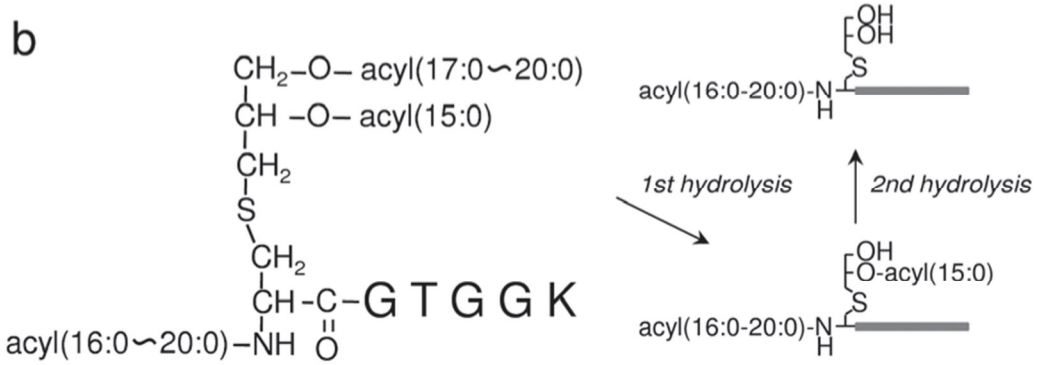


Fig. 7 *N*-acylated triacyl structure of *S. aureus* lipoprotein SitC. (a) MALDI-TOF MS analysis pattern of SitC fragments, which were obtained by lysylendopeptidase digestion in an SDS-polyacrylamide gel slice. A series of mass peaks exhibiting 14 mass differences are numbered 1–8 in the left panel and listed in the right panel. The obtained *m/z* of peak 1 corresponded to the theoretical *m/z* of tripalmitic acid (Pam_3)-modified *N*-acyl-*S*-(diacyl-propyl)-cysteinyl-peptide, and those of peaks 2–8 corresponded to those of Pam_3-N-terminal peptides harbouring increasing numbers of methylene (CH2) groups in their fatty acids. (b) Treatment of triacylated SitC with lipoprotein lipase (LPL). LPL specifically hydrolyses ester-bonded fatty acids; triacylated N-terminal SitC lipopeptides in the left panel were hydrolysed into diacylated and then monoacylated forms, as indicated on the right. The triacylated N-terminal SitC structure was determined by combined MS/MS analysis after treatment with lipoprotein lipase. The exact positions and lengths of ester-linked fatty acids could not be determined, but were predicted from additional information on the lipoprotein structures of *Bacillus* species (Kurokawa *et al.*, 2012) and the phospholipid structures of *S. aureus* (Fischer, 1994).

To confirm that the α-amino group of the *S*-diacyl-glyceryl-Cys residue is acylated, a peak corresponding to the octadecanoyl-glyceryl lipopeptide of SitC generated after long term treatment with LPL was further analysed by MALDI-ion trap MS/MS. The results strongly supported that the fatty acid modification site is the N-terminal Cys residue and that the octadecanoyl group is linked to the α-amino group of Cys via an amide bond. These results indicate that the N-terminal Cys of SitC from *S. aureus* RN4220 cells is *N*-acylated with a saturated C16 to C20 fatty acid and modified with a diacylglyceryl group containing two saturated fatty acids (Fig. 7b, left). Triacylation of SitC was confirmed with *S. aureus* SA113, clinically isolated *S. aureus* MW2, MSSA476, and *S. epidermidis* cells (Asanuma *et al.*, 2011). Furthermore, other *S. aureus* lipoproteins such as SA0739, SA0771, SA2074 (ModA), SA2158 and SA2202 were found to be triacylated. Taken together, these results indicate that staphylococcal species, typical monoderm bacteria, synthesize *N*-acylated *S*-diacylglyceryl lipoproteins even though an *E. coli* Lnt homologue is absent.

N-terminal lipopeptide structures of other monoderms

The structures of N-terminal lipopeptides generated from the lipoproteins of other low GC monoderms were analysed. Novel structures found in these bacteria (Fig. 8a) were *N*-acyl-*S*-monoacyl-glyceryl-Cys (lyso form) in *Enterococcus faecalis*, *N*-acetyl-*S*-diacyl-glyceryl-Cys (*N*-acetyl form) in *Bacillus licheniformis* and dipeptidyl-*S*-diacyl-glyceryl-Cys (peptidyl form) in *Mycoplasma fermentans* (Kurokawa et al., 2012b).

The unexpected lyso-form structure reflects the loss of one of two esterified fatty acids linked to the glyceryl group (Fig. 8a). Three lipoproteins of *E. faecalis*, an intestinal bacterium, were found to be of the lyso-form (Kurokawa *et al.*, 2012b) (Table 1). In addition, lyso-form lipoproteins were found in *Bacillus cereus*, a component of the microbiota in the human gut and also a food poisoning bacterium, *Lactobacillus delbrueckii* subsp. *bulgaricus*, a probiotic strain originating from Bulgarian yogurt, and *Streptococcus sanguinis*, a member of the human indigenous oral microflora. These results suggest that lyso-form lipoproteins are one of the common structures in low GC content Gram-positive bacteria. It might be possible that lyso-form lipoproteins function as immune regulatory molecules in the gut or play a role in the host gut mucosal immunity.

In addition to *Bacillus licheniformis*, a soil bacterium causing food poisoning, an *N*-acetyl form was found in *Bacillus subtilis* and strains living in extreme environments such as alkalophilic *Bacillus halodurans*, alkalophilic and extremely halotolerant *Oceanobacillus*

iheyensis, and thermophilic *Geobacillus kaustophilus* (Kurokawa et al., 2012b). Therefore, both *N*-acetyl-and lyso-forms are found in the *Bacillaceae* family (Table 1).

The MBIO_0661 lipoprotein of *Mycoplasma fermentans* was found to be a dipeptidyl form comprising Ala-Gly (Kurokawa *et al.*, 2012b). In contrast, the dipeptide of the MBIO_0319 lipoprotein was Ala-Ser, while the MBIO_0763 and MBIO_0869 lipoproteins were the conventional diacyl form (Table 1). These extension sequences coincided with the C-terminal sequences of the respective signal peptides, suggesting that *M. fermentans* Lsp has unusual substrate specificity. Lipoproteins purified from *M. genitalium* and *M. pneumoniae* were found to be triacylated (Kurokawa *et al.*, 2012b), which was consistent with previous results obtained in immunological studies (Shimizu *et al.*, 2007; Shimizu *et al.*, 2008). These results indicate that lipoprotein structures are diverse in mycoplasma species and vary depending on even lipoprotein species in *M. fermentans*.

It should be noted that three lipoproteins of *Listeria monocytogenes* were found to be of the conventional diacyl form (Kurokawa *et al.*, 2012b).

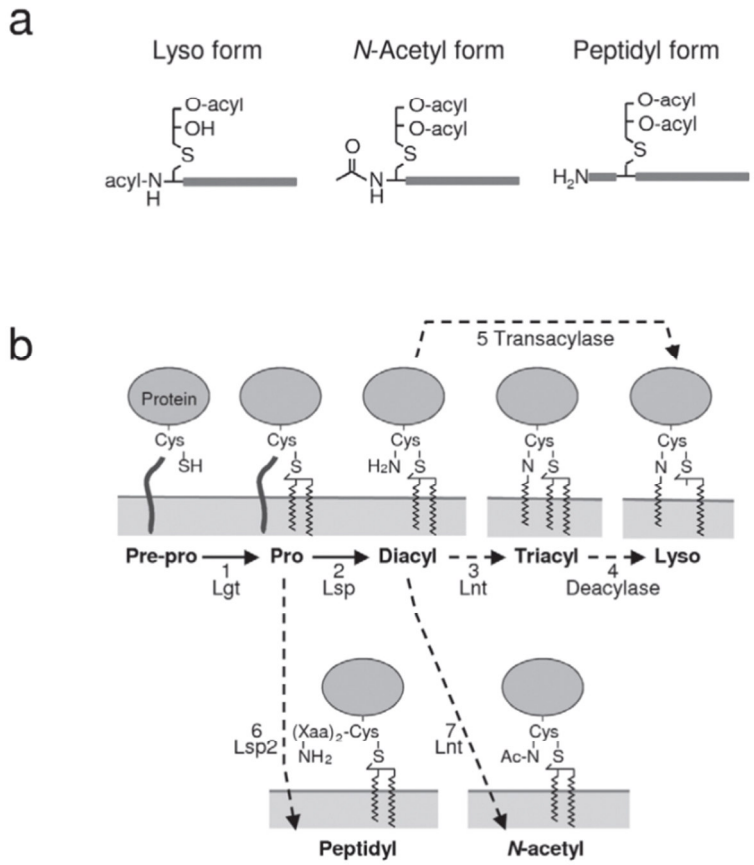


Fig. 8 Novel *N*-terminal structures of lipoproteins and their possible biosynthetic pathways. (a) The structures of the lyso-form, *N*-acetyl form, and peptidyl form are illustrated. (b) Putative lipoprotein biogenesis pathways in low GC content monoderms are shown. The dotted lines indicate uncharacterized steps.

Putative lipoprotein biogenesis pathways

Based on the results of recent studies (Asanuma *et al.*, 2011; Kurokawa *et al.*, 2009; Kurokawa *et al.*, 2012b; Serebryakova *et al.*, 2011), the N-terminal structures of lipoproteins in monoderm bacteria and mycoplasma species are classified into three major classes (A to C) (Table 1). Possible biosynthetic pathways are shown in Fig. 8b. Class C lipoproteins having *S*-diacylglyceryl-Cys are generated by two enzymes, Lgt and Lsp (steps 1 and 2 or 6). A novel type of Lsp (Lsp 2) is postulated, which cleaves inside the signal peptide, probably after a serine residue, and leaves a peptide preceding the lipidated Cys (step 6). Class B lipoproteins having *N*-acyl-*S*-diacyl-glyceryl-Cys can be generated through sequential modification reactions by Lgt, Lsp and an unidentified Lnt (steps 1, 2, and 3 or 7). Since an *E. coli*-type Lnt

Table 1. The biochemical characterization of recently identified bacterial lipoproteins

		Modified groups to the lipidated cysteine					
Bacterial species	Major habitat	Protein name	Total[1]	R1[1]	R2[2]	R3[3]	Reference
Class A. *N*-acyl-*S*-monoacylglyceryl-cysteine structures							
A-a. Lyso form							
B. cereus	Intestine	BC0200	32:0	17:0	H	15:0	Kurokawa *et al.* (2012a)
	Foods[3]	OppA	32:0	17:0, 18:0	H	14:0, 15:0	Kurokawa *et al.* (2012a)
		PrsA	32:0	17:0	H	15:0	Kurokawa *et al.* (2012a)
E. faecalis	Intestine	EF2256	34:1	18:1	H	16:0	Kurokawa *et al.* (2012a)
		EF3256	34:1	18:1	H	16:0	Kurokawa *et al.* (2012a)
		PnrA	34:1	18:0, 18:1	H	16:0, 16:1	Kurokawa *et al.* (2012a)
L. bulgaricus	Intestine	Ldb0202		18:1	H	18:1	Kurokawa *et al.* (2012a)
		Ldb2183		18:1	H	18:1	Kurokawa *et al.* (2012a)
S. sanguinis	Oral cavity	SSA_0375		18:1	H	16:0	Kurokawa *et al.* (2012a)
		SSA_1038		18:0, 18:1	H	16:0, 16:1	Kurokawa *et al.* (2012a)

Class B. *N*-acyl-*S*-diacylglyceryl-cysteine structures							
B-a. N-*acetyl form*							
B. halodurans	Soil	BH3460	32:0	15:0	15:0	2:0	Kurokawa *et al.* (2012b)
		MalE	32:0	15:0	15:0	2:0	Kurokawa *et al.* (2012b)
B. licheniformis	Soil	MntA	34:0	17:0	15:0	2:0	Kurokawa *et al.* (2012b)
		OppA	34:0	17:0	15:0	2:0	Kurokawa *et al.* (2012b)
B. subtilis	Soil	BSU01630	34:0	17:0	15:0	2:0	Kurokawa *et al.* (2012b)
		PrsA	34:0	17:0	15:0	2:0	Kurokawa *et al.* (2012b)
G. kaustophilus	Deep sea	GK1283	34:0	32:04		2:0	Kurokawa *et al.* (2012b)
			34:0	17:0	15:0	2:0	Kurokawa *et al.* (2012b)
O. iheyensis	Deep sea	CtaC	32:0	15:0	15:0	2:0	Kurokawa *et al.* (2012b)
B-b. Conventional triacyl form							
A. laidlawii	Waste water	ACL_1223	48:0, 50:0[5]	16:0	16:0	16:0, 18:0	Serebryakova *et al.* (2011)
			50:0	(18:0)[6]	(16:0)[6]	(16:0)[6]	Nakayama *et al.* (2012)
		ACL_1410	48:0, 50:0[5]	16:0	16:0	16:0, 18:0	Serebryakova *et al.* (2011)
			50:0	(18:0)[6]	(16:0)[6]	(16:0)[6]	Nakayama *et al.* (2012)
M .genitalium	Genital	MG_040	50:1	18:1	16:0	16:0	Kurokawa *et al.* (2012b)
M. pneumonia	Respira-tory tracts	MPN052	50:1	18:1	16:0	16:0	Kurokawa *et al.* (2012b)
		MPN415	50:1	18:1	16:0	16:0	Kurokawa *et al.* (2012b)
S. aureus	Nares, skin	SA0739	52:0	32:0–37:0[4]		15:0–20:0	Asanuma *et al.* (2011)
		SA0771	52:0	32:0–36:0[4]		16:0–20:0	Asanuma *et al.* (2011)

		SA2074	51:0	31:0–36:0[4]		15:0–20:0	Asanuma *et al.* (2011)
		SA2202	52:0	32:0–36:0[4]		16:0–20:0	Asanuma *et al.* (2011)
		SitC	53:0	33:0–37:0[4]		16:0–20:0	Asanuma *et al.* (2011)
S. epidermidis	Skin	SitC	53:0	32:0–35:0[4]		17:0–20:0	Kurokawa *et al.* (2012b)
Class C. *S*-diacylglyceryl-cysteine structures							
C-a. Conventional diacyl form							
L. monocytogenes	Foods[3]	Lmo0135	32:0	17:0	15:0	H	Kurokawa *et al.* (2012b)
		Lmo2196	32:0	17:0	15:0	H	Kurokawa *et al.* (2012b)
		Lmo2219	32:0	17:0	15:0	H	Kurokawa *et al.* (2012b)
M. fermentans	Throat	MBIO_0763	34:0	34:0[4]		H	Kurokawa *et al.* (2012b)
		MBIO_0869	34:0	34:0[4]		H	Kurokawa *et al.* (2012b)
S. aureus	Nares, skin	SA1659	33:0	18:0	15:0	H	Kurokawa *et al.* (2012a)
		SitC	32:0–35:0	17:0–20:0	15:0	H	Kurokawa *et al.* (2012a)
C-b. Peptidyl form							
M. fermentans	Throat	MBIO_0319	34:0	18:0	16:0	Ala-Ser-	Kurokawa *et al.* (2012b)
		MBIO_0661	34:0	18:0	16:0	Ala-Gly-	Kurokawa *et al.* (2012b)

[1] The total carbon number of modified acyl groups on the lipidated cysteine is calculated from the MS/MS-analyzed lipopeptide peak. The most abundant N-terminal lipopeptide ion in MALDI-TOF MS was usually analyzed by MS/MS.
[2] R1 and R2 denote a hydrogen or an acyl group attached to the *sn*-1 and *sn*-2 positions of the *S*-glyceryl group of the lipidated cysteine, respectively. R3 denotes a hydrogen, an acyl group, or a dipeptide attached to the α-amino group of the lipidated cysteine.
[3] Foods probably contained human, animal, and/or soil contaminants.
[4] The total carbon number of modified acyl groups on R1 and R2 is shown.
[5] Descending order in abundance.
[6] Our interpretation is shown in parenthesis.

homologue is not found in monoderm bacteria, not only the structure but also the properties of this putative Lnt seem to be significantly different from those of *E. coli*-type Lnt. Class A lipoproteins having *N*-acyl-*S*-monoacyl-glyceryl-Cys can be generated by either a putative

lipoprotein *O*-deacylase from the triacyl form (step 4) or a putative lipoprotein transacylase from diacyl form (step 5). In *E. coli*, *N*-acylation is essential for correct outer membrane localization of lipoproteins by the Lol system (Fukuda *et al.*, 2002; Robichon *et al.*, 2005). To understand the biological significance of lipoproteins in low GC content monoderms, it is important to characterize the enzymes involved in lipoprotein modification.

It seems interesting to consider the fatty acid preference of the putative Lnt enzymes involved in the formation of the triacyl form in monoderms (Kurokawa *et al.*, 2012a). In *S. aureus*, 48% of fatty acids bound to the α-amino group of Cys (R3 position) are 18:0 (18 carbons with no double bond), whereas 18:0 fatty acid constitutes only 3.2% of total membrane lipids and 13.4% of PG. In contrast, 15:0 and 20:0 fatty acids occupy 40% and 22% of total membrane fatty acids, respectively. These results suggest that Lnt of *S. aureus* seems to prefer 18:0 fatty acid although the source of the fatty acid remains to be determined. On the other hand, the R3 position of the *M. pneumoniae* MPN052 lipoprotein carried only 16:0 fatty acid, while *M. pneumoniae* total membrane contained 18:0 and 18:1 fatty acids, each at about 20%, in addition to 16:0. Thus, the putative Lnt of *M. pneumonia* is specific to 16:0 fatty acid. Moreover, lyso form lipoproteins such as PrsA of *B. cereus*, PnrA of *E. faecalis*, and Ldb2183 of *L. bulgaricus* also exhibit strong fatty acid specificity at the R3 position (Kurokawa *et al.*, 2012b). Lnt of *E. coli* was recently found to exhibit specificities for both fatty acid species and lipid head groups (Hillmann *et al.*, 2011). Kurokawa *et al*. (2012a) discovered conditions that alter between triacyl and diacyl forms of lipoproteins in *S. aureus*. Under the combination of low pH and post-log growth phase, *S. aureus* cells accumulated N-terminal free diacyl forms of lipoproteins. High temperatures and/or high salt concentrations additively increased the accumulation of diacyl forms. Interestingly, pH shift assays revealed that protein synthesis is required to the structural alterations, suggesting that the expression level or activity of unidentified Lnt should be regulated via protein synthesis.

Biological functions of monoderm-derived lipoproteins

Mammalian TLRs play an important role in recognizing microorganisms to activate host innate immune responses (Takeuchi & Akira, 2010). Eleven human TLRs and 13 mouse TLRs have been identified, and each TLR appears to recognize a pathogen-associated molecular pattern molecule derived from various microorganisms, including bacteria, viruses, protozoa, and fungi (Takeuchi & Akira, 2010). Among them, TLR4, TLR5 and TLR9 recognize a single class of pattern molecule, such as lipopolysaccharides (LPS), bacterial

flagellin, and bacterial DNA, respectively (Ewald & Barton, 2011; Ramos *et al.*, 2004). Outer membrane LPS play a critical role in the activation of TLR4 in diderm Gram-negative bacteria (Hoshino *et al.*, 1999). TLR2 plays a major role in recognition of Gram-positive bacteria (Takeuchi *et al.*, 1999). TLR2 had been reported to recognize several molecules, including lipoproteins, synthetic lipopeptides, peptidoglycan, lipoteichoic acids (LTAs), lipomannans and lipoarabinomannans (Zahringer *et al.*, 2008). Since these molecules are structurally diverse, it appeared unlikely that TLR2 has the ability to react with all agonists to the same degree. To address this issue, two mutant strains of *S. aureus*, Δ*ltaS* (Grundling & Schneewind, 2007; Oku et al., 2009) and Δ*lgt* (Stoll et al., 2005), were used. It was found that bacterial lipoproteins, but neither LTA nor peptidoglycan, act as native TLR2 ligands (Kurokawa *et al.*, 2009). The *lgt*-deficient mutants of *S. aureus*, *L. monocytogenes*, and Group B *Streptococcus* also indicated that bacterial lipoproteins function as major ligands for TLR2 (Bubeck Wardenburg *et al.*, 2006; Hashimoto *et al.*, 2006; Henneke *et al.*, 2008; Machata *et al.*, 2008). Some lipoproteins tightly associate with peptidoglycan. When peptidoglycan-associated lipoproteins were enzymatically removed, TLR2 activation by peptidoglycan was completely abolished, while lipoproteins extracted from peptidoglycan remained active as to TLR2 activation (Kurokawa *et al.*, 2009). Taken together, these results indicate that true TLR2 ligands are lipoproteins.

Because *S. aureus* lipoproteins anchored to the cytoplasmic membrane are covered by a thick layer of peptidoglycan (Fig. 1), bacterial cells must be engulfed and delivered to acidic phagosomes for the efficient activation of TLR2. Enzymes in phagosomes digest the internalized bacteria, causing the release of bacterial lipoproteins. Thus, phagocytosis of *S. aureus* cells and following degradation of the cell wall are critical for an efficient lipoprotein–TLR2 interaction (Ip *et al.*, 2010; Kang *et al.*, 2011; Shimada *et al.*, 2010).

TLR2 was thought to function as a heterodimer with TLR1 or TLR6. It was then examined as to which heterodimer is involved in the cytokine production by triacylated SitC. Synthetic lipoprotein analogues, such as *N*-palmitoyl-*S*-dipalmitoylglyceryl (Pam_3)-Cys-Ser-Lys-Lys-Lys-Lys (Pam_3CSK_4) and *S*-dipalmitoylated macrophage-activating lipopeptide-2 kDa (MALP-2) (Muhlradt *et al.*, 1997), were used to examine the pro-inflammatory cytokine release. These examinations suggested that the TLR2/TLR1 heterodimer generally recognizes triacylated lipopeptides whereas the TLR2/TLR6 heterodimer responds to diacylated lipopeptides (Buwitt-Beckmann *et al.*, 2005; Omueti *et al.*, 2005; Takeuchi *et al.*, 2001; Takeuchi *et al.*, 2002). The crystal structures of the two TLR2 heterodimers complexed with synthetic lipopeptide support this model (Jin *et al.*, 2007; Kang *et al.*, 2009).

Induction of pro-inflammatory cytokines by purified triacylated SitC was examined using mouse thioglycolate-elicited peritoneal macrophages. The native SitC protein induced the production of both tumour necrosis factor-α (TNF-α) and interleukin 6 (IL-6) in wild-type mouse macrophages (Fig. 9) (Kurokawa *et al.*, 2009). In contrast, TLR2–/– mouse macrophages released neither TNF-α nor IL-6 upon the addition of SitC. Since both cytokines were induced in macrophages from TLR1–/– and TLR6–/– mice, the two TLRs are not essential for SitC-mediated cytokine production while TLR2 is essential, suggesting that triacylated SitC stimulates immune cells through both the TLR2/TLR1 and TLR2/ TLR6 heterodimers.

Lyso form lipoprotein OppA of *B. cereus* induced both TNF-α and IL-6 in a TLR2- dependent manner (Fig. 9) (Kurokawa *et al.*, 2012b). Neither TLR1 nor TLR6 is required for this. In contrast, cytokine secretion induced by lyso form lipoprotein PnrA of *E. faecalis* is dependent on both TLR2 and TLR6 but independent of TLR1 (Fig. 9). *N*-acetyl lipoprotein/lipopeptide induced the release of TNF-α and IL-6 in TLR2- and TLR6-dependent and TLR1-independent manners (Fig. 9). Taken together, these results indicate that the lyso and *N*-acetyl form lipoproteins in monoderms activate TLR2 by forming a heterodimer with either TLR1 or TLR6. The choice of co-receptor seems to depend on the species of bacterial lipoproteins.

PROBLEMS TO BE ANSWERED

As discussed in this chapter, much new information has been obtained in relation to the N-terminal structures of lipoproteins, the enzymes involved in lipoprotein maturation, the physiological functions of lipoproteins, and the sorting of lipoproteins. However, there are many unanswered questions, as listed below.

Gram-negative diderms

1 The Lol avoidance mechanism is probably general for the cytoplasmic membrane retention of lipoproteins. However, how do different residues function as Lol avoidance signals when the origin of LolCDE is different?

2 How are lipoproteins transported to the cell surface of some bacteria? Does the Lol system contribute to this transport?

Gram-positive diderms

1 How are lipoproteins transported to the mycobacterial outer membrane?

2 Which lipoprotein motifs direct retention in the cytoplasmic membrane or transport to the mycobacterial outer membrane?
3 What is the function of lipoprotein glycosylation in mycobacteria?
4 Is the glycosylation of lipoproteins dependent on lipidation?

Low GC monoderms

1 Why is the lipid modification of N-terminal Cys so diverse in low-GC monoderms?
2 Are modification enzymes also diverse?
3 Is there any physiological advantage of synthesizing *N*-acylated triacyl lipoproteins even though bacteria do not possess an outer membrane?
4 Is structural alteration between lipid modifications unique in *S. aureus*, or common in other bacteria?

ACKNOWLEDGEMENTS

HT was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan. PS was supported by the Swiss National Science Foundation (SNF: 31003A_135705). LBL was supported by a Korean Research Foundation Grant (KRF: 2010–0020665).

ADDENDUM

Personal contribution to chapter 1

My contribution as a co-author to this manuscript was as follows:

- Database- and literature search about mycobacterial lipoproteins
- Writing the part of the manuscript about mycobacterial lipoproteins

CHAPTER 2

CLONING, EXPRESSION AND CHARACTERIZATION OF *MYCOBACTERIUM TUBERCULOSIS* LIPOPROTEIN LPRF

Juliane K. Brülle*, Thomas Grau*, Andreas Tschumi, Yolanda Auchli, Reto Burri, Silke Polsfuss, Peter M. Keller, Peter Hunziker, Peter Sander

Biochemical and Biophysical Research Communications 391 (2010) 679–684

ABSTRACT

Lipoproteins are well known virulence factors of bacterial pathogens in general and of *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis, in particular. Lipoprotein lipidation between Gram-positive and Gram-negative bacteria differs significantly as these are di- and triacylated, respectively. Little is known about the lipid anchor of mycobacterial lipoproteins. We reported recently that mycobacterial LppX, a lipoprotein involved in synthesis of cell wall components is triacylated, although mycobacteria are classified as GC-rich Gram-positive bacteria. We here exploited the model organism *Mycobacterium smegmatis* for the expression of *Mtb* LprF and characterized N-terminal modifications at the molecular level. LprF is a putative lipoprotein of *Mtb* involved in signaling of potassium-dependent osmotic stress. LprF is extensively modified in a mycobacterium-specific manner by a thioether-linked diacylglyceryl residue with one ester-bound tuberculostearic- and one C16:0 fatty acid and additionally by a third *N*-linked C16:0 fatty acid, and a hexose.

* These authors contributed equally to this work

INTRODUCTION

Tuberculosis is a major cause of death around the world, with 9.3 million new cases and 1.8 million deaths occurring in 2007, which is the highest rate claimed by a single bacterial pathogen (WHO TB-factsheet, 2009). The causative agent of the disease is *Mycobacterium tuberculosis (Mtb)*, an acid-fast bacillus that is primarily transmitted via the respiratory route. The reasons for the pathogens extraordinary success are diverse: it is slow-growing which makes antibiotic treatment complicated and lengthy, it has a thick waxy cell wall and therefore is resistant to different kinds of mechanical and chemical stress and it evades the immune system by parasitizing the macrophages of its host.

The high immunogenic potential of *Mtb* is based on its unusual cell envelope which is exceptionally rich in lipids, glycolipids and polysaccharides (Daffe & Draper, 1998). Among others, uncommon cell wall components like mycolic acids, mycocerosic acid, phenolthiocerol, lipoarabinomannan (LAM) and arabinogalactan trigger inflammatory host reactions (Daffe & Draper, 1998). On the other hand, *Mtb* is able to subvert the immune response of the host by inhibiting its innate defense by several mechanisms. It prevents inflammasome activation (Master *et al.*, 2008), delays phagosome maturation in macrophages (Russell, 2001) and suppresses MHC class II antigen presentation (Baena & Porcelli, 2009). Suppression of MHC II antigen expression is conferred by several lipoproteins, e.g., the 19 kDa lipoprotein (LpqH) (Noss *et al.*, 2001).

Lipoproteins are a subclass of proteins found in the cell envelope of all bacteria. Lipoproteins are either di- (in case of Gram-positive bacteria) or triacylated (in case of Gram-negative bacteria) on a highly conserved cysteine located at the N-terminus, which is part of the lipobox [LVI][ASTVI][GAS][C] (Rezwan *et al.*, 2007a). By the consecutive action of the three enzymes pre-prolipoprotein diacyl glyceryl transferase (Lgt), prolipoprotein signal peptidase (LspA) and apolipoprotein *N*-acyltransferase (Lnt), lipoproteins are post-translationally modified after translocation over the cytoplasmic membrane. Lgt attaches a diacylglyceryl residue to the universally conserved cysteine in the lipobox by thioether-linkage. Then LspA removes the lipoprotein signal peptide and Lnt attaches a third acyl chain to the amino group of the modified cysteine. Despite the fact that *N*-acylation in *Bacillus subtilis* and *Staphylococcus aureus* was reported (Hayashi *et al.*, 1985; Kurokawa *et al.*, 2009; Tawaratsumida *et al.*, 2009), Lnt was found exclusively in Gram-negative bacteria. Even though *Mtb* is deemed to be a Gram-positive bacterium because of its staining properties, it has been shown recently that mycobacteria have a periplasmic-like structure (Zuber *et al.*,

2008). Furthermore, it has been shown that mycobacteria express a functional Lnt as at least one lipoprotein, LppX was found to be triacylated (Tschumi *et al.*, 2009).

The functions of lipoproteins are manifold; they may be involved in protein export and folding, in antibiotic resistance, in ABC transporter systems, act as substrate-binding proteins and are involved in cell signaling. Examples of mycobacterial lipoproteins which have been explored are Mpt83, LppX, LpqW and the 19 kDa lipoprotein. Mpt83 is assumed to be an adhesin and it has been shown that RNA encoding Mpt83 induces protective immune responses against *Mtb* infection (Xue *et al.*, 2004). LpqW and LppX have been shown to be key players in synthesis and transport of unique components of the mycobacterial cell envelope. While LppX is involved in translocation of phthiocerol dimycocerosates (DIM) to the outer membrane (Sulzenbacher *et al.*, 2006), LpqW has been shown to be essential in the synthesis of the cell wall components phosphatidyl-*myo*-inositol mannoside (PIM) and LAM (Kovacevic *et al.*, 2006; Marland *et al.*, 2006). The 19 kDa lipoprotein has been described as an adhesin (Diaz-Silvestre *et al.*, 2005). It induces IL-1, IL-2 and TNF-α through TLR2-signaling in macrophages. Overall, lipoproteins are important in host–pathogen interactions and they have a high pathogenic potential which has been proven by disruption of *lspA*, the lipoprotein signal peptidase. An *lspA*-deficient strain of *Mtb* exhibited reduced multiplication in mouse macrophages and reduced number of colony forming units in a mouse model by 3–4 logs (Sander *et al.*, 2004).

The putative lipoproteins LprF and LprJ of *Mtb* recently have been described to interact with the histidine kinase KdpD in a yeast two-hybrid screen (Steyn *et al.*, 2003). Both lipoproteins have been suggested to form ternary complexes with the histidine kinase domain of KdpD which in turn seems to be activated after potassium-dependent sensing of environmental osmotic stress and activates a signal transduction pathway.

Even though advances in the past few years contributed to the knowledge of the function of lipoproteins, there is scarcely known anything about the chemical composition of the lipid modifications. Modern tools like mass spectrometry instead of incorporation of radioactive precursors provide support to understand post-translational modifications by lipidation and glycosylation of lipoproteins. Mycobacterial lipoprotein LppX is the first and only lipoprotein characterized at the molecular level (Tschumi *et al.*, 2009). Investigations in other bacteria indicate that lipids of lipoproteins may differ significantly within one species (Kurokawa *et al.*, 2009; Tawaratsumida *et al.*, 2009). Therefore, we extended our studies on mycobacterial lipoproteins.

In this study, we cloned, expressed and purified LprF of *Mtb* in *Mycobacterium smegmatis*, a non-pathogenic but phylogenetically close relative of *Mtb*. We show for the first time that a putatively membrane-localized lipoprotein of *Mtb* is *N*-acylated and determine the predominant fatty acids by matrix-assisted laser desorption ionization time-of-flight tandem mass spectrometry (MALDI-TOF/TOF MS).

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. *Mycobacterium smegmatis* strains were cultivated in LB medium or on LB agar. If necessary, gentamicin was added at a final concentration of 5 μg/ml. Strains used in this study were *M. smegmatis* SmR5 (Sander *et al.*, 1995), a strain carrying a non-restrictive *rpsL* mutation conferring streptomycin resistance, and *M. smegmatis lnt::aph* mutant (Tschumi *et al.*, 2009).

Expression of recombinant lipoprotein LprF. LprF of *Mtb* was expressed under control of the 19 kDa antigen (*lpqH*) promoter. The recombinant protein was generated by fusion PCR; fragment 1 contains the sequence of the 19 kDa promoter, fragment 2 contains the functional domain of the LprF lipoprotein including the leader peptide and the sequence of a C-terminal thrombin cleavage site and a HA-epitope tag (hemagglutinin protein), followed by a 6× His-epitope tag. Fragment 1 was amplified with primers 1a (5′-GG**GTTAACGAATTC** TACATTG-3′, bold letters indicate sites of restriction enzymes *Hpa*I and *Eco*RI) and 1b (5′-*CTTGTGAGATCAAGCCATTCAT*CCTGTGCTCCT-3′, italic letters indicate the linker for *lprF*) from plasmid pMV261-Gm-FusLppX (Tschumi *et al.*, 2009). Fragment 2 was amplified with primers 2a (5′-AGGAGCACAGG*ATGAATGGCTTGATCTCACAAG*-3′, italic letters indicate the linker for *lprF*) and 2b (5′-CC**GTTAACGAATTC**TAGTGGTGGTG GTGGTGGTGAGTGGCGTAGTCGGGGACGTCGTAGGGGTAACTACCACGTGGAAC TAGTCCCGCCGGGTTCGG-3′, bold letters indicate sites of restriction enzymes *Hpa*I and *Eco*RI, underlined letters indicate thrombin cleavage site, HA- and His-tag) from *Mtb* genomic DNA. Fusion PCR was performed with primers 1a and 2b. The recombinant *lprF* gene was cloned into pMV261-Gm (a shuttle vector replicating in *Escherichia coli* as well as in mycobacteria, derived from pMV261) using the *Eco*RI-restriction sites. The resulting plasmid pMV261-Gm-*lprF* was transformed into *M. smegmatis* SmR5 and *M. smegmatis lnt::aph*.

Preparation of cell extracts and Western blot analysis. Bacteria transformed with pMV261-Gm-*lprF* were cultured in 2 L LB medium for 3 days at 37 C°. The cultures were harvested (4400 rpm, 1 h) and resuspended in PBS containing Complete EDTA-free tablets (Roche). Cells were lysed by three French press cycles (American Instrument Company) at 1.1×10^6 Pa. Extracts were treated with 2% sodium *N*-lauroylsarcosine (SLS) for 1 h at room temperature, and for 16 h at 4 C° thereafter. Soluble and insoluble fractions were separated by centrifugation at 30,000*g* for 1 h at 4 C°. Extracts corresponding to 1–5 µg of total protein were separated by a 12.5% SDS–PAGE and subsequently analyzed by Western blot using anti-HA-antibodies (1:300, Roche).

Protein fingerprinting. Proteins were digested with trypsin and dissolved in 25 µl 0.1% formic acid. Samples were desalted by using a Ziptip C18 column, mixed 1:1 with matrix solution (5 mg/ml α-cyano-4-hydroxy-cinnamic acid (Bruker Daltonics) in 0.1% trifluoroacetic acid, 50% acetonitrile) and spotted onto the target.

FPLC protein purification. The soluble fraction of cell extracts expressing epitope-tagged proteins was mixed with dilution buffer (20 mM NaH_2PO_4, 0.5 M NaCl, pH 7.4) and loaded on a HisTrap™ HP column (GE Healthcare) previously equilibrated with equilibration buffer (20 mM NaH_2PO_4, 0.5 M NaCl, 0.2% SLS, 20 mM imidazole, pH 7.4). Proteins were eluted applying an imidazole gradient (0.125–0.5 M).
As a further purification step, the column flow through from His-tag purification was dialyzed against equilibration buffer (20 mM Tris-hydroxymethyl-aminomethane, 0.1 M NaCl, 0.1 mM EDTA, pH 7.5), loaded onto an anti-HA-affinity matrix (Roche), washed with a buffer containing 20 mM Tris-hydroxymethyl-aminomethane, 0.1 M NaCl, 0.1 mM EDTA, 0.05% (v/v) Tween 20, pH 7.5 and eluted with column regeneration buffer (0.1 M glycine, pH 2.0).

MALDI-TOF/TOF. 100–200 pmol of purified lipoprotein were prepared and analyzed according to Ujihara et al. (Ujihara *et al.*, 2008). Lipoproteins were digested with AspN. Extracted peptides were dried and dissolved in 5 µl 0.1% trifluoroacetic acid, 50% acetonitrile. Samples were mixed 1:1 with matrix solution (5 mg/ml α-cyano-4-hydroxy-cinnamic acid (Bruker Daltonics) in 0.1% trifluoroacetic acid, 50% acetonitrile) and spotted onto the target. The MALDI-TOF/TOF mass spectra were recorded on an Ultraflex II MALDI-TOF/TOF instrument with smartbeam laser upgrade (Bruker Daltonics). The laser was set to a repetition rate of 100 Hz and the ion acceleration voltage was 29.5 kV. The mass measurements were performed in the positive ion reflector mode.

Edman degradation. For N-terminal sequencing, proteins were separated by 12.5% SDS–PAGE, blotted onto a PVDF membrane and stained with Coomassie Brilliant Blue (0.03% (w/v) Coomassie Brilliant Blue R-250, 25% (v/v) isopropanol, 10% (v/v) acetic acid) for 5 min. The membrane was destained (10% acetic acid, 35% methanol) and washed once with Aqua B. Braun (Ecotainer®). Visible bands were cut out and analyzed by Edman degradation (Procise 492 cLC protein sequencer, Applied Biosystems), according to the manufacturer's instructions.

RESULTS AND DISCUSSION

To analyze the putative lipoprotein LprF from *Mtb* concerning the lipoprotein specific modifications with fatty acids we generated the expression vector pMV261-Gm-*lprF*. Plasmid pMV261-Gm-*lprF* was transformed into *M. smegmatis* SmR5 wildtype and an isogenic *lnt*::*aph* mutant strain, lacking a functional apolipoprotein-*N*-acyl-transferase (Lnt). Recombinant LprF from whole cell extracts was purified using the His-epitope and subsequently analyzed by Western blot using anti-HA-antibodies. Depending on the kinetics of the enzymes of the lipoprotein biosynthesis pathway (Rezwan *et al.*, 2007a), the cell extract may contain different forms of the lipoprotein, the pre-prolipoprotein, prolipoprotein, apolipoprotein and the mature lipoprotein. The theoretically calculated molecular masses for LprF are in the range of 26–30 kDa. However, the apparent molecular masses estimated from an SDS–PAGE may differ significantly. Analysis of the elution fractions from wildtype and *lnt::aph* mutant by Western blot and Coomassie stained SDS–PAGE showed bands with an apparent size of 30 and 35 kDa in both strains (data not shown).

Protein analysis

Fingerprint analysis of the 30 and 35 kDa proteins from wildtype and *lnt::aph* mutant confirmed these proteins as LprF from *Mtb*. For characterization and identification of the N-terminal modifications of LprF only the mature lipoprotein is needed. To characterize the 30 and 35 kDa forms of LprF which have been detected after His-tag purification we applied several methods. Edman degradation provides a suitable method to identify the pre-prolipoprotein, prolipoprotein and apolipoprotein. *N*-acylation blocks Edman degradation.

Therefore, MALDI-TOF/TOF MS was used to confirm *N*-acylated LprF forms and to identify modifications at the molecular level.

Edman degradation

The proteins with an apparent size of 35 kDa (Fig. 1A) revealed the sequence MNGLI, which is an LprF sequence starting with the initial methionine of the signal peptide (Fig. 1C) thereby confirming these forms as pre-pro-LprF or pro-LprF in both strains. The proteins with a size of 30 kDa (Fig. 1A) revealed the sequence KKPTT in wildtype and TVVAG in the *lnt::aph* mutant. These determined sequences start at position +3 and at position −5, respectively, in relation to the cysteine (+1) of the lipobox (Fig. 1C).

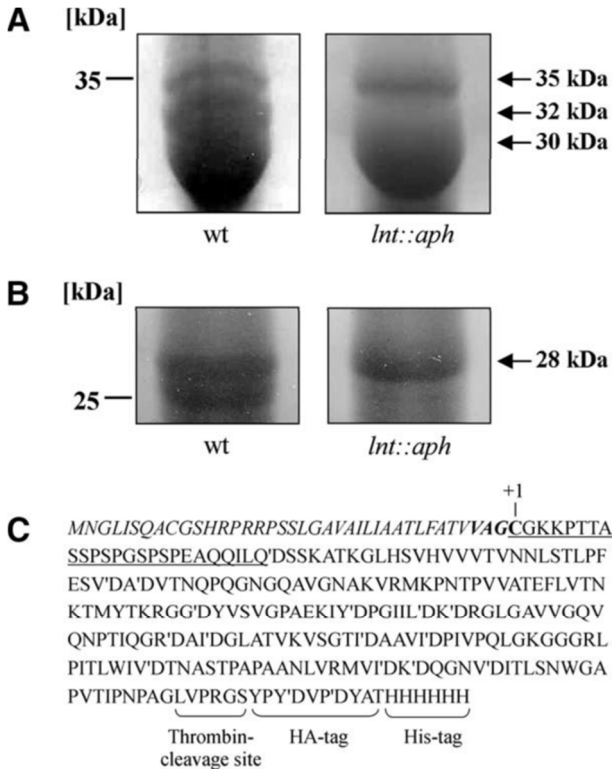


Fig. 1. Copper chloride stained SDS–PAGE and amino acid sequence of purified LprF. (A) HisTrap™-column purified LprF from *M. smegmatis* wildtype and *lnt::aph* mutant. (B) HA-affinity matrix purified LprF from *M. smegmatis* wildtype and *lnt::aph* mutant. Proteins at 25 kDa were identified as histone-like protein HupB from *M. smegmatis*. (C) Amino acid sequence of recombinant LprF. Italic letters indicate the signal peptide cleaved by LspA. Bold letters indicate the lipobox including the conserved cysteine at position +1 modified by Lgt and Lnt. Inverted commas indicate AspN cleavage sites. The modified N-terminal peptide after AspN digestion found in *M. smegmatis* wt and *lnt::aph* mutant is underlined.

These results indicate that the 30 kDa forms of LprF are not the desired LspA-cleaved LprF but LprF forms cleaved by other proteases. The 30 kDa protein form isolated from the *lnt::aph* mutant and starting with the sequence TVVAG contains potentially diacylglycerol

modified cysteine and therefore was subjected to MALDI-TOF/TOF MS analysis. The 30 kDa band isolated from the wildtype strain may be a mixture of triacylated LprF and truncated LprF with the N-terminal residues KKPTT. As triacylated lipoprotein is not accessible to Edman degradation due to the blocked N-terminus, the 30 kDa band from wildtype was also subjected to MALDI-TOF/TOF MS.

MALDI-TOF/TOF MS analysis

The copper chloride stained SDS–PAGE of His-tag-purified LprF showed the same bands detected before on Western blot and Coomassie stained SDS–PAGE (Fig. 1A). AspN-digested peptides from the 30 kDa recombinant LprF of both strains were analyzed with MALDI-TOF/TOF MS to confirm the N-terminal sequence of mature LprF and to characterize the predominant modifications occurring in *M. smegmatis* lipoprotein LprF at the molecular level. The calculated monoisotopic *m/z* value for the AspN digested unmodified N-terminal peptide of the LspA-cleaved LprF is *m/z* = 2496.2 (Fig. 1C). The fatty acids found in mycobacterial phospholipids are palmitic (C16:0), palmitoleic (C16:1), oleic (C18:1) and tuberculostearic acid (10-methyloctadecanoic acid) (C19:0) (Goren, 1979). Since fatty acids of membrane phospholipids are used for *N*-acylation of lipoproteins in *E. coli* (Jackowski & Rock, 1986; Lai & Wu, 1980), we calculated the theoretical mass of the N-terminal peptide of LprF with all possible combinations of the above mentioned four fatty acids. Lipoproteins sometimes are glycosylated (Sartain & Belisle, 2009) and putative sites for *O*-glycosylation are also present in the N-terminal AspN-fragment of LprF. Therefore, we also calculated the mass with several hexose modifications. However, no signals corresponding to the free or acylated N-terminal peptides (with or without glycosylations) were found in the mass spectrum (data not shown). So, the analyzed 30 kDa proteins of LprF are most likely the non-modified but truncated proteins, as indicated by the results of the Edman degradation.

Analysis of HA-tag purified LprF

Since the 35 kDa LprF clearly was identified as the (pre-) pro-LprF and the 30 kDa forms were not the LspA-processed LprF we assumed that the expected mature forms were not yet isolated. Concentrating again on the Western blots and Coomassie stained SDS–PAGE from His-tag purification, we found another distinct band on the Western blots from the column

flow throughs. This band had an apparent size of 28 kDa possibly corresponding to the LspA-cleaved LprF (data not shown).
Due to the fact that this protein form did not bind to the HisTrap™ HP column, we used the HA-epitope in the recombinant LprF to purify this protein from the flow through fractions. After purification with the HA-affinity matrix, the proteins with the size of 28 kDa in wildtype and *lnt::aph* mutant were observed on Western blot, Coomassie and copper chloride stained SDS–PAGE as well (Fig. 1B).

Edman degradation of HA-tag purified LprF

To determine whether these 28 kDa forms from wildtype and *lnt::aph* mutant are the LspA-cleaved LprF or LprF at all, Edman degradation was performed although for the mature lipoprotein from wildtype N-terminal sequencing is expected to be blocked by the modified N-terminus. The analyses revealed the unique sequence xxKKP in LprF indicating that these proteins start with the cysteine at position +1 (Fig. 1C) thereby confirming these proteins as the LspA-cleaved LprF in the wildtype and the mutant. As the resolution of the SDS–PAGE is not sufficient to separate triacylated and diacylated lipoproteins, the 28 kDa band from wildtype may be a mixture of both forms.

MALDI-TOF/TOF MS analysis of HA-tag purified LprF

For the identification of the lipoprotein specific modifications at the conserved cysteine the AspN digested peptides from the 28 kDa LprF from wildtype and *lnt::aph* mutant were investigated by MALDI-TOF MS. Instead of the $[M+H]^+$ signal at m/z = 2496.2 as calculated for the AspN digested, unmodified N-terminal peptide we found a signal at m/z = 3488.9 for the modified N-terminal peptide of LprF from wildtype (Fig. 2A). In the *lnt::aph* mutant we found a signal at m/z = 3250.5 (Fig. 2B), indicating a smaller size of the N-terminal peptide and thus suggesting an Lnt dependent modification in the wildtype. The difference in molecular mass between the unmodified N-terminal peptide (m/z = 2496.2) and the peptide found in the *lnt::aph* mutant (m/z = 3250.5) is 754.3 Da indicating a diacylglyceryl modification with ester-linked tuberculostearic acid and C16:0 fatty acid (592.5 Da) and a glycosylation with one hexose (162.2 Da, Σ = 754.7). The difference in molecular mass of 238.4 Da between wildtype (m/z = 3488.9) and mutant indicates an additional modification of the N-terminal peptide with a C16:0 fatty acid in the wildtype.

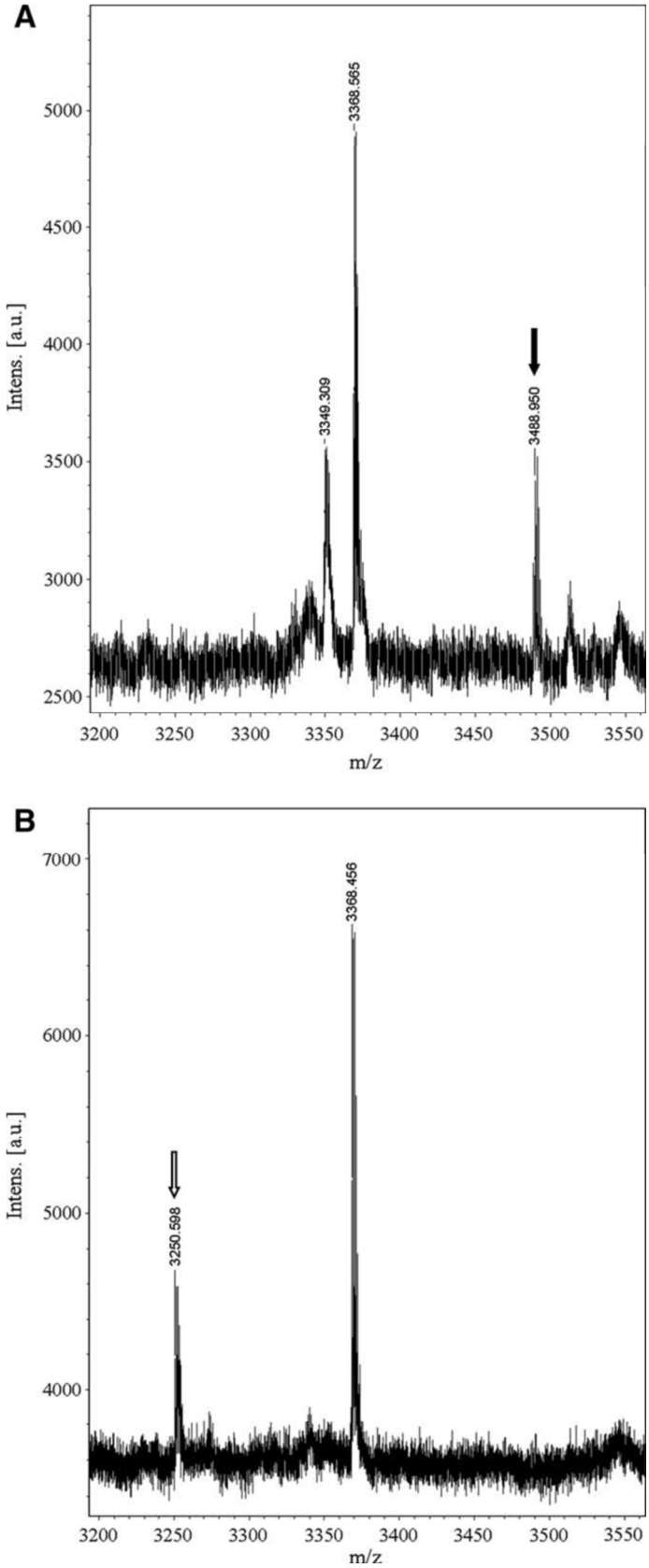


Fig. 2. MALDI-TOF MS analysis of AspN digested purified LprF (28 kDa). MS data of AspN digested LprF peptides from (A) *M. smegmatis* wildtype and (B) *M. smegmatislnt::aph* mutant. Black arrow indicates triacylated monoglycosylated N-terminal peptide. Open arrow indicates diacylated monoglycosylated N-terminal peptide. The mass at *m/z* = 3368.5 corresponds to an internal AspN peptide of LprF (residue 201–232 of the mature lipoprotein).

In order to confirm the modifications of the AspN digested N-terminal peptide of LprF from wildtype (*m/z* = 3488.9) and *lnt::aph* mutant (*m/z* = 3250.5), the structure of the N-terminal peptide was analyzed with MALDI-TOF/TOF MS. In wildtype (Fig. 3A) the ion at *m/z* = 3326.7 corresponds to the cleavage of a hexose (*Δ* = 161.2 Da). The ion at *m/z* = 2698.7 is the most intense ion and corresponds to the release of the diacylthioglyceryl carrying both, an *O*-linked tuberculostearic and a C16:0 fatty acid (*Δ* = 628.0 Da). The release of 257.1 Da from the ion at *m/z* = 3326.7 corresponds to the elimination of a C16:0 fatty acid and the release of 370.9 Da from the ion at *m/z* = 3069.6 corresponds to the elimination of a tuberculostearic acid α-thioglyceryl ester. The difference in molecular mass between the MS signal for the N-terminal peptide from wildtype and *lnt::aph* mutant is 238.4 Da which indicates a third acylation with a C16:0 fatty acid in the wildtype peptide in an Lnt dependent manner.
In the *lnt::aph* mutant (Fig. 3B) the ions at *m/z* = 3088.8 and 2461.9 correspond to the release of a hexose (*Δ* = 161.8) and a diacylthioglyceryl carrying both *O*-linked tuberculostearic and C16:0 fatty acid (*Δ* = 626.9), respectively. Of note – a loss of the His-epitope in this 28 kDa LprF forms was assumed to be responsible for the failure to extract this proteins by the His-tag purification. But, the presence of the tag was confirmed by MALDI-TOF/TOF MS analysis (data not shown).

CONCLUSIONS

This study shows for the first time the modifications of a putative membrane located lipoprotein of *Mtb* on the molecular level. The LprF of *Mtb* wildtype is a triacylated and glycosylated lipoprotein carrying a thioether-linked diacylglyceryl residue with an ester-bound tuberculostearic- and C16:0 fatty acid and a third C16:0 fatty acid most likely at the amino terminal cysteine residue. These results together with the recent analysis of LppX (Marland *et al.*, 2006) indicate that *N*-acylation seems to be a common motif in mycobacterial lipoproteins. As expected for the *lnt::aph* mutant the cysteine in LprF is modified with the diacylglyceryl residue, but is missing the third acylation (only glycosylated and modified with the diacylglyceryl residue), thereby confirming lipoprotein acylation by MSMEG_3860 in mycobacteria. Due to the close phylogenetic relation between *M. smegmatis* and *Mtb* the same acylation pattern of LprF in *Mtb* is assumed. Although glycosylation of the N-terminal AspN peptide of LprF was identified, the exact glycosylation site of seven possible sites within the peptide could not be determined.

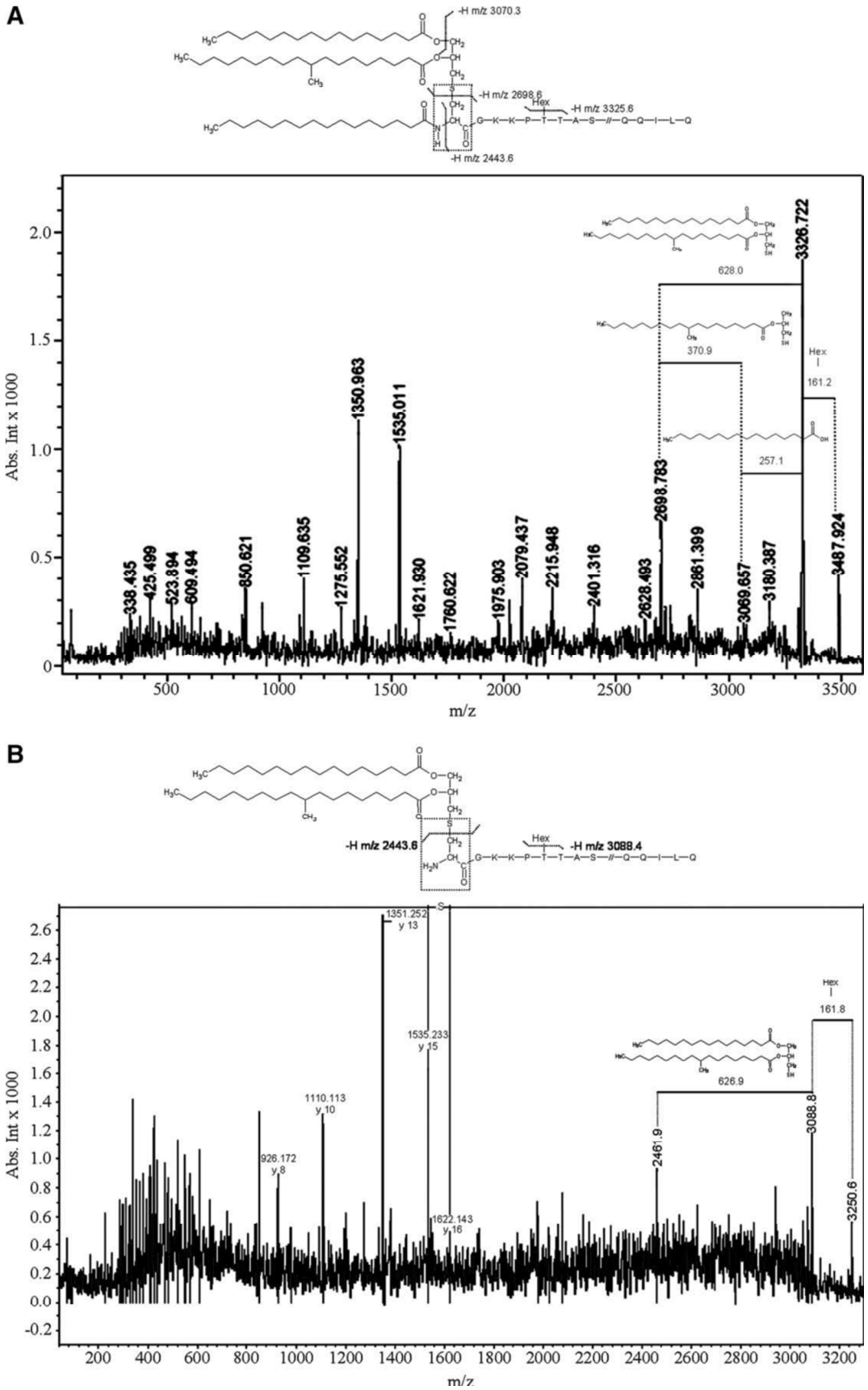


Fig. 3. MALDI-TOF/TOF MS analysis of N-terminal peptides of LprF. MS/MS data of the N-terminal peptides of LprF from (A) *M. smegmatis* wildtype and (B) *M. smegmatis lnt::aph* mutant. Schematic drawings of the structure of the modified N-terminal peptide of LprF are shown in the upper part of each MS/MS spectrum. Dotted frames indicate the modified conserved cysteine at molecular level. Cleavage sites of each identified *m/z* signal are indicated. Eliminated fragments of LprF modifications are depicted in the spectra.

Fingerprinting and Edman degradation identified several proteins isolated either by His- or HA-affinity purification as LprF. However, the N-terminal lipidated LprF could not be purified by His-tag purification suggesting that the lipidation interferes with His-purification despite the presence of the His-tag which was proved by analysis of the C-terminus (data not shown) or somehow reduces the binding affinity of the protein. Beside the pre-pro-lipoprotein, apolipoprotein and the mature LprF a protein with an apparent size of 32 kDa found only in the wildtype was confirmed by fingerprint as LprF as well, but was not subjected to further analysis in this study. This LprF may result from cleavage by proteases other than LspA.

ACKNOWLEDGMENTS

We thank Rene Brunisholz from the Functional Genomics Center Zurich for performing N-terminal sequencing. This work was supported in part by the Swiss National Science Foundation (3100A0-120326) and the European Union (TB-Drug: LSHP-CT-2006-037217).

ADDENDUM

Personal contribution to chapter 2

My contribution as shared-first author to this manuscript was as follows:

- Design of the study
- Generation of expression vectors
- Lipoprotein purification
- MALDI-TOF data analysis
- MALDI-TOF/TOF data analysis
- Writing of the manuscript

CHAPTER 3

LIPOPROTEINS OF SLOW-GROWING MYCOBACTERIA CARRY THREE FATTY ACIDS AND ARE *N*-ACYLATED BY APOLIPOPROTEIN *N*-ACYLTRANSFERASE BCG_2070C

Juliane K. Brülle, Andreas Tschumi, and Peter Sander

BMC Microbiology 2013, 13:223

ABSTRACT

Background Lipoproteins are virulence factors of *Mycobacterium tuberculosis*. Bacterial lipoproteins are modified by the consecutive action of preprolipoprotein diacylglyceryl transferase (Lgt), prolipoprotein signal peptidase (LspA) and apolipoprotein *N*-acyltransferase (Lnt) leading to the formation of mature triacylated lipoproteins. Lnt homologues are found in Gram-negative and high GC-rich Gram-positive, but not in low GC-rich Gram-positive bacteria, although *N*-acylation is observed. In fast-growing *Mycobacterium smegmatis*, the molecular structure of the lipid modification of lipoproteins was resolved recently as a diacylglyceryl residue carrying ester-bound palmitic acid and ester-bound tuberculostearic acid and an additional amide-bound palmitic acid.

Results We exploit the vaccine strain *Mycobacterium bovis* BCG as model organism to investigate lipoprotein modifications in slow-growing mycobacteria. Using *Escherichia coli* Lnt as a query in BLASTp search, we identified BCG_2070c and BCG_2279c as putative *lnt* genes in *M. bovis* BCG. Lipoproteins LprF, LpqH, LpqL and LppX were expressed in *M. bovis* BCG and BCG_2070c *lnt* knock-out mutant and lipid modifications were analyzed at molecular level by matrix-assisted laser desorption ionization time-of-flight/time-of-flight analysis. Lipoprotein *N*-acylation was observed in wildtype but not in BCG_2070c mutants. Lipoprotein *N*- acylation with palmitoyl and tuberculostearyl residues was observed.

Conclusions Lipoproteins are triacylated in slow-growing mycobacteria. BCG_2070c encodes a functional Lnt in *M. bovis* BCG. We identified mycobacteria-specific tuberculostearic acid as further substrate for *N*-acylation in slow-growing mycobacteria.

INTRODUCTION

Proteins posttranslationally modified by covalent lipid attachment are present in eukaryal and bacterial organisms. In bacteria, 1-3 % of the genome encode for lipoproteins. Bacterial lipoproteins are anchored in the membrane surface where they fulfill various cellular functions, ranging from cell wall integrity, secretion, nutrient uptake, environmental signaling to virulence (Babu *et al.*, 2006; Kovacs-Simon *et al.*, 2011; Sutcliffe & Harrington, 2004).

Lipoproteins from Gram-positive and Gram-negative bacteria are synthesized in the cytosol as preprolipoproteins and are translocated across the cytoplasmic membrane via the Sec- or Twin arginine translocation (Tat) system (McDonough *et al.*, 2005). Lipoprotein signal sequences terminate in a highly conserved lipobox motif consisting of four amino acids (LVI/ASTVI/GAS/C) (Babu *et al.*, 2006). Processing of lipoprotein precursors into mature forms takes place at the outer leaflet of the cytoplasmic membrane and is accomplished by the sequential action of three enzymes attacking the conserved cysteine in the lipobox: 1) the phosphatidylglycerol:pre-prolipoprotein diacylglyceryl transferase (Lgt) attaches a diacylglyceryl residue to the cysteine via thioether linkage (Sankaran & Wu, 1994), 2) the prolipoprotein signal peptidase (LspA) cleaves off the signal peptide and 3) apolipoprotein *N*-acyltransferase (Lnt) acylates the N-terminal cysteine residue at its free amino group (Okuda & Tokuda, 2011; Rezwan *et al.*, 2007a; Sutcliffe & Harrington, 2004). In proteobacteria, *N*-acylation of lipoproteins is a prerequisite for the transport to the outer membrane by the Lol system (Narita & Tokuda, 2011; Yakushi *et al.*, 2000).

Lgt and LspA are universally present in Gram-positive and Gram-negative bacteria (Wu, 1996). The gene encoding Lnt was originally identified in the Gram-negative bacterium *Salmonella enterica* sv. Typhimurium and is conserved in proteobacteria. The Lnt structure and function are well studied in *Escherichia coli* (Vidal-Ingigliardi *et al.*, 2007). Contrary to the long held assumption that *lnt* is restricted to Gram-negative bacteria (Wu, 1996) *lnt* homologues are also present in high GC-rich Gram-positive bacteria. In the fast-growing, saprophytic mycobacterial model organism *Mycobacterium smegmatis*, Lnt-dependent *N*-acylation was demonstrated and the lipid moiety of lipoproteins has been resolved at molecular level. *M. smegmatis* lipoproteins are modified with a thioether-linked diacylglyceryl residue composed of ester-linked palmitic acid and ester-linked tuberculostearic acid and an additional palmitic acid amide-linked to the α-amino group of the conserved cysteine. Diacylglycerol modification and signal peptide cleavage are prerequisites for *N*-acylation (Brülle *et al.*, 2010; Tschumi *et al.*, 2009). Secreted proteins, among them

lipoproteins often are modified by glycosylation. *O*-glycosylation in mycobacteria occurs through a stepwise process depending on at least a protein mannosyl tranferase (PMT) performing the initial mannosylation step and a α1-2 mannosyl tranferase realizing the subsequent elongation of the mannosyl chains. Recently, PMT enzyme responsible for the initial attachment of mannose residue to the protein was identified (Liu *et al.*, 2013).

In addition to *M. smegmatis*, *N*-acyltransferase activity by Lnt homologues was shown in two other high GC-rich Gram-positive bacteria, namely *Streptomyces scabies* (Widdick *et al.*, 2011) and *Corynebacterium glutamicum* (Mohiman *et al.*, 2012). Recent mass spectrometry analyses of lipoproteins in low GC-rich Gram-positive bacteria (firmicutes and mollicutes) provided evidence that *N*-acylation also occurs in these bacterial species, however, no obvious *lnt*-like gene has been identified to date (Hayashi *et al.*, 1985; Kurokawa *et al.*, 2009; Nakayama *et al.*, 2012; Serebryakova *et al.*, 2011; Tawaratsumida *et al.*, 2009). Instead, biochemical analysis identified two new lipoprotein structures, the "*N*-acetyl" and the "peptidyl" lipoprotein structure. These novel structures strongly suggest the presence of yet to be identified key enzymes involved in bacterial lipoprotein biosynthesis (Kurokawa *et al.*, 2012b).

Most pathogenic mycobacteria belong to the group of slow-growing mycobacteria, including *Mycobacterium leprae*, the causative agent of leprosy and the members of the *Mycobacterium tuberculosis* complex (e.g. *M. tuberculosis, Mycobacterium africanum, Mycobacterium cannetti, Mycobacterium bovis*). *Mycobacterium tuberculosis* is the causative agent of human tuberculosis, a major cause of death around the world (http://www.who.int/tb/publications/factsheets/en/index.html). Elimination of tuberculosis requires an improved understanding of the host, the pathogen and their interaction for the development of better, more effective drugs and vaccines. Lipoprotein biogenesis is a major virulence factor of *M. tuberculosis* (Rampini *et al.*, 2008; Sander *et al.*, 2004). Moreover, lipoproteins evidently meet pathogen-associated molecular patterns (PAMPs) criteria and are well detected by innate immune recognition mechanisms (Ray *et al.*, 2013). *M. tuberculosis* lipoproteins are major antigens and trigger the activation of cellular and humoral immune responses to mycobacteria. Lipoproteins are potent agonists of toll-like receptor 2 (TLR2) which upon long term stimulation has been associated with the down regulation or deviation of the immune response. TLR2 agonist activity has been demonstrated for several *M. tuberculosis* lipoproteins including LpqH, LprA, LprG and PstSI (Drage *et al.*, 2009; Harding & Boom, 2010). Recently, it was reported that mycobacteria generate and release membrane vesicles (MVs) (Prados-Rosales *et al.*, 2013). Strikingly, MVs from pathogenic mycobacteria as

compared to non-pathogenic mycobacteria are enriched in lipoproteins, some of them well known TLR2 agonists. MVs produced a severe TLR2 dependent inflammatory response in vitro and in vivo (Prados-Rosales *et al.*, 2013). Investigations regarding the vaccine potential of MVs from pathogenic mycobacteria elicited a mixed cellular and humoral immune response. This suggests a vaccine potential of MVs and their lipoproteins against *M. tuberculosis*.

Even though research on lipoproteins in fast-growing mycobacteria contributed to the knowledge of lipoprotein biosynthesis and modification, there is scarcely known anything about lipoprotein modifications and their chemical structures in slow-growing mycobacteria. *Mycobacterium bovis* bacille Calmette Guerin (BCG) is derived from virulent *M. bovis*, the causative agent of bovine tuberculosis. The genome of *M. bovis* BCG is highly similar to the *M. tuberculosis* genome (>99.5% sequence identity) (Brosch *et al.*, 2007). *M. bovis* BCG was first used in 1921 as a live vaccine against tuberculosis. Since then four billion doses have been applied to humans. Still today it is the only licensed tuberculosis vaccine, despite its incomplete protective efficacy, particular against adult lung tuberculosis (Kaufmann & Gengenbacher, 2012).

Concerning the presence of open reading frames (ORFs) encoding lipoprotein modifying enzymes, both genomes of *M. tuberculosis* and *M. bovis* BCG Pasteur reveal a single ORF encoding Lgt (Rv1614, BCG_1652) and a single ORF encoding LspA (Rv1539, BCG_1591). Two ORFs encoding Lnt are found in *M. bovis* BCG (BCG_2070c, BCG_2279c). BCG_2070c (which is identical to *M. tuberculosis* Rv2051c = *ppm1*) is a two domain protein with a conserved apolipoprotein-*N*-acyltransferase and a Ppm-like domain. BCG_2279c shows conserved apolipoprotein-*N*-acyltransferase domain and exhibits considerable homology to *E. coli* Lnt. In *M. tuberculosis*, the corresponding open reading frame is split into two, Rv2262c and Rv2261c. In our previous analysis (Tschumi *et al.*, 2009), these may have escaped our attention, since split. Only upon completion of the *M. bov*is BCG sequence the homology to Lnt became apparent. Due to this polymorphism in the second *M. tuberculosis* putative Lnt ORF, we focussed our studies on lipoproteins and lipoprotein synthesis in slow-growing mycobacteria on the vaccine strain *M. bovis* BCG. Prediction of lipoproteins in *M. tuberculosis* complex using DOLOP database suggests the presence of 50 potential lipoproteins of the approximately 4000 ORFs (Babu *et al.*, 2006). However, the existence of twice as many lipoproteins has been discussed (Sutcliffe & Harrington, 2004).

In this study, we show that lipoproteins are triacylated in slow-growing *M. bovis* BCG. We demonstrate apolipoprotein *N*-acyltransferase acitivity and by targeted gene deletion identify

BCG_2070c as a functional Lnt. We give structural information about the lipid modification of four mycobacterial lipoproteins, LprF, LpqH, LpqL and LppX. Hereby mycobacteria-specific tuberculostearic acid is identified as a further substrate for *N*-acylation.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions. Mycobacterium bovis BCG Pasteur strains were cultivated in Middlebrook 7H9 medium or on Middlebrook 7H10 agar enriched with oleic acid albumin dextrose (OADC, Difco). Liquid broth was supplemented with 0.05 % of Tween 80 to avoid clumping. If necessary, the appropriate antibiotic was added at the following concentration: 5 μg ml^{-1} gentamicin, 100 μg ml^{-1} streptomycin, 25 μg ml^{-1} hygromycin. Strains used in this study were *M. bovis* BCG SmR (further referred to as *M. bovis* BCG or parental strain) (Sander *et al.*, 2001), a streptomycin resistant derivative of *M. bovis* BCG Pasteur 1173P2, Δ*lnt* = *M. bovis* BCG SmR *lnt* knock out mutant in BCG_2070c and Δ*lnt*-*lnt*BCG_2070c = *M. bovis* BCG SmR *lnt* knock out mutant in BCG_2070c transformed with complementing vector pMV361-hyg-*lnt*BCG_2070c.

Disruption of lnt *in* M. bovis *BCG.* A 1.9 kbp *Mlu*I/*Nsi*I fragment of *M. bovis* BCG from position 2296156 to 2294306 comprising the 5'*lnt* flanking sequence and a 2.8 kbp *SnaB*I/*Mlu*I fragment from position 2292652 to 2289856 comprising the 3'*lnt* flanking sequence of the *lnt* domain of BCG_2070c were PCR amplified using genomic DNA from *M. bovis* BCG Pasteur and cloned into vector pMCS5-rpsL-hyg with the respective enzymes resulting in knock-out vector pMCS5-rpsL-hyg-Δ*lnt*BCG. This way, we deleted a 1.6 kbp of the Lnt domain without introducing a frameshift or any other deletion to the Ppm synthase domain. The *lnt*BCG allele was deleted in the *M. bovis* BCG SmR chromosome as described previously (Sander *et al.*, 1995; Sander *et al.*, 2001) and confirmed by Southern blot analysis with 0.2 kbp *Sal*I *lnt* downstream probe. For complementation with *M. bovis* BCG BCG_2070c a 6.3 kbp fragment from *M. bovis* BCG from position 2289839 to 2296178 spanning the entire *lnt* gene was cloned into pGEM-T Easy (Promega) to result in pGEM-T Easy-*lnt*BCG_2070c and subsequently subcloned as a 6.3 kbp *Eco*RI fragment into the *Hpa*I site of plasmid pMV361-hyg (Sander *et al.*, 1997) to result in pMV361-hyg- *lnt*BCG_2070c. Complementation was confirmed by Southern blot analyses with 0.2 kbp *Kpn*I/*Hind*III *lnt*BCG_2070c upstream probe.

Expression of Lipoproteins LprF, LpqH, LpqL and LppX. Plasmid pMV261-Gm, a derivative of pMV261 shuttle vector, is able to replicate in *E. coli* as well as in mycobacteria (Stover *et al.*, 1991). *LprF* (Brülle *et al.*, 2010), *lpqH*, *lpqL* and *lppX* (Tschumi *et al.*, 2009) were amplified by PCR from *M. tuberculosis* genomic DNA and fused to the *M. tuberculosis* 19 kDa promoter. The target proteins and 19 kDa promoter are identical between *M. tuberculosis* and *M. bovis* BCG. Sequences encoding a hemagglutinin and a hexa- histidine epitope were fused to the 3' part of each gene to facilitate subsequent purification and detection on Western blot. The insert was cloned into the *EcoR*I site of pMV261-Gm to result in pMV261-Gm-LprF, pMV261-Gm-LpqH, pMV261-Gm-LpqL and pMV261-Gm-LppX. Subsequently plasmids were transformed into BCG parental strain, Δ*lnt* and Δ*lnt*-*lnt*BCG_2070c.

Preparation of cell extracts and Western blot analysis. Bacteria from 1-liter cultures were harvested and resuspended in phosphate-buffered saline containing Complete EDTA-free tablets (Roche) to inhibit protein degradation. Cells were lysed by three French Press cycles (American Instrument Co.) at 1.1 x 10^6 Pa. Extracts were treated with 2% sodium *N*-lauroylsarcosine (SLS) for 1 h at room temperature, and incubated for 16 h at 4 °C thereafter. Extracts corresponding to 1-5 µg of total protein were separated by a 12.5% SDS-PAGE gel and subsequently analyzed by Western blot using anti-HA-antibody (1:300, Roche) and corresponding secondary antibody conjugated with horseradish peroxidase.

Fast Protein Liquid Chromatography Protein Purification. Soluble fractions of cell extracts from recombinant strains expressing epitope-tagged proteins were diluted with buffer containing 20 mM NaH_2PO_4, 0.5 M NaCl, pH 7.4 to 1% sodium *N*-lauroylsarcosine and loaded on a HisTrap™ HP column (GE Healthcare) previously equilibrated with buffer containing 20 mM NaH_2PO_4, 0.5 M NaCl, 0.2% sodium *N*-lauroylsarcosine and 20 mM imidazole, pH 7.4. Proteins were eluted applying an imidazole gradient (0.125-0.5 M). As a further purification step, if necessary, HisTrap™ HP column flow through was dialyzed against buffer containing 20 mM Tris-hydroxymethyl-aminomethane, 0.1 M NaCl, 0.1 mM EDTA, pH 7.5 and loaded onto anti-HA-affinity matrix (Roche). Proteins were eluted with buffer containing 0.1 M glycine, pH 2.0.

MALDI-TOF/TOF analysis. 100-200 pmol of purified lipoprotein were prepared and analyzed according to Ujihara *et al.* (Ujihara *et al.*, 2008). Briefly, lipoproteins in elution fractions from

FPLC or HA chromatography were precipitated and SDS-PAGE gel was performed. Proteins separated by electrophoresis were visualized with copper staining. Protein bands with the apparent molecular weight of apolipoprotein/mature lipoprotein were cut from the stained gel. Lipoproteins were in-gel digested with Trypsin or AspN and extracted peptides were dried and dissolved in 5 µl 0.1% trifluoroacetic acid, 50% acetonitrile. Samples were loaded onto the target and covered with 1 µl matrix solution (5 mg ml^{-1} α-cyano-4-hydroxy-cinnamic acid (Bruker Daltonics) in 0.1% trifluoroacetic acid, 50% acetonitrile). The MALDI-TOF/TOF mass spectra were recorded on an Ultraflex II MALDI-TOF/TOF instrument with smartbeam laser upgrade (Bruker Daltonics). The laser was set to a repetition rate of 100 Hz and the ion acceleration voltage was 29.5 kV. The mass measurements were performed in the positive ion reflector mode.

RESULTS

Lipoproteins are expressed in M. bovis *BCG*

As model substrates for lipoprotein modification in slow-growing mycobacteria we chose four different lipoproteins being identical in *M. tuberculosis* and in *M. bovis* BCG Pasteur. The well characterized LppX (Sulzenbacher *et al.*, 2006; Tschumi *et al.*, 2009) and LprF (Brülle *et al.*, 2010) in addition to LpqH and LpqL. LppX (Rv2945c) has been shown to be involved in translocation of phthiocerol dimycocerosates (DIM) to the outer membrane (Sulzenbacher *et al.*, 2006). LprF (Rv1368) is involved in signaling and has been suggested to interact with the histidine kinase KdpD in response to environmental osmotic stress (Steyn *et al.*, 2003). LpqH (19 kDa antigen, Rv3763) functions as an adhesin and has been recognized as an immunodominant lipoprotein (Diaz-Silvestre *et al.*, 2005). LpqL (Rv0418) is predicted to be a lipoprotein aminopeptidase. Hence, our choice of lipoproteins is representing different classes of lipoproteins. The four expression vectors pMV261-Gm for hexa-histidine/hemagglutinine tagged LprF, LpqH, LpqL or LppX were transformed into *M. bovis* BCG. Whole cell extracts from the four strains expressing the recombinant lipoproteins were analyzed by Western blot. The apparent molecular masses of the detected proteins correspond to the predicted mass of the recombinant apolipoproteins/mature lipoproteins (LprF 29.4 kDa, LpqH 17.3 kDa, LpqL 54.2 kDa, LppX 26.3 kDa). Eventually the prepro-/pro-lipoprotein forms whose sizes are increased by 2-3 kDa due to the presence of the signal peptide, are also detected.

Identification of the lipoprotein lipid anchor in M. bovis *BCG*

To characterize the modifications of lipoproteins at the molecular level, the four recombinant lipoproteins LprF, LpqH, LpqL and LppX were expressed in *M. bovis* BCG parental strain. Proteins were purified by FPLC or HA affinity chromatography. Eluted fractions were analyzed by Western blot (see additional file 2) and lipoprotein containing fractions were precipitated for SDS-PAGE gel. Bands of purified lipoproteins were visualized with copper staining, cut from the gel and the proteins were in-gel digested with Trypsin or AspN (in case of LprF). Resulting peptides were prepared and analyzed by MALDI-TOF/TOF mass spectrometry (Ujihara *et al.*, 2008). For the identification of the modification we determined the structure and calculated the expected monoisotopic molecular masses of the unmodified N-terminal tryptic or AspN-digested peptides of LprF, LpqH, LpqL and LppX (without signal peptide). Phospholipids found in mycobacteria mainly consist of palmitic (C16:0), palmitoleic (C16:1), oleic (C18:1) and tuberculostearic acid (10-methyloctadecanoic acid) (C19:0) (Goren, 1979). In *E. coli*, fatty acids of membrane phospholipids, i.e. myristic (C14:0), palmitic, palmitoleic, oleic (C18:1 ω9) or vaccenic (18:1 ω7) acid are used for the modification of lipoproteins (Gupta *et al.*, 1991; Hillmann *et al.*, 2011; Jackowski & Rock, 1986; Lai & Wu, 1980; Lin *et al.*, 1980). Therefore we calculated the theoretical mass of the N-terminal peptides of the four lipoproteins with all possible combinations of the above mentioned fatty acids observed in mycobacterial phospholipids to identify putative modifications. Glycosylations are also commonly found in lipoproteins (Garbe *et al.*, 1993; Sartain & Belisle, 2009). Some of the analyzed N-terminal peptides carry putative *O*-glycosylation sites, therefore we also calculated the masses with hexose modifications. $[M+H]^+$ signals at *m/z* values which we calculated for the unmodified N-terminal peptides were not found. Instead, we found MS signals at *m/z* values which indicate that the N-terminal peptides are modified in a lipoprotein-specific manner with different combinations of saturated and unsaturated C16, C18 and C19 fatty acids. The calculated *m/z* values are summarized and compared with the experimentally determined *m/z* values in Table 1.

The modifications we estimated from the $[M+H]^+$ signals in the MS spectrum were confirmed by MS/MS fragmentation and thereby information about the linkage of the modification was obtained. The structures of the di- or triacylated N-terminal tryptic or AspN-digested peptides from LprF, LpqH, LpqL and LppX were investigated by MS/MS. All eliminations found in MS/MS of lipoproteins isolated from the parental strain are summarized in Table 2.

TABLE 1. Comparison of *m/z* values of N-terminal AspN-digested/tryptic peptides of LprF, LpqH, LpqL and LppX found in BCG parental and Δ*lnt* mutant strain.

	Peptide	Calculated *m/z*	Parental strain *m/z*	*Δlnt* *m/z*
LprF	**CGK…ILQ**	**2496.24**	-	-
	CGK…ILQ + Diacylglycerol (C16/C16)	3047.11 (+550.87)	-	3046.70 (+550.46)
	CGK…ILQ + Diacylglycerol (C16/C18)	3073.15 (+576.91)	-	3072.71 (+576.47)
	CGK…ILQ + Diacylglycerol (C16/C19)	3089.20 (+592.96)	-	3088.74 (+592.50)
	CGK…ILQ + Diacylglycerol (C16/C19) + Hexose	3251.44 (+755.20)	-	3251.65 (+755.41)
	CGK…ILQ + Diacylglycerol (C16/C19) + *N*-acyl (C16)	3327.60 (+831.36)	3326.83 (+830.59)	-
	CGK…ILQ + Diacylglycerol (C16/C19) + *N*-acyl (C19) + Hexose	3531.93 (+1035.69)	3530.56 (+1034.32)	-
LpqH	**CSSNK**	**538.23**	-	-
	CSSNK + Diacylglycerol (C16/C16)	1089.10 (+550.87)	-	1088.60 (+550.37)
	CSSNK + Diacylglycerol (C16/C18)	1115.14 (+576.91)	-	1114.68 (+576.45)
	CSSNK + Diacylglycerol (C16/C19)	1131.19 (+592.96)	1130.79 (+592.56)	1130.71 (+592.48)
	CSSNK + Diacylglycerol (C16/C19) + *N*-acyl (C16)	1369.59 (+831.36)	1369.04 (+830.81)	-
LpqL	**CIR**	**391.21**	-	-
	CIR + Diacylglycerol (C16/C19)	984.17 (+592.96)	984.50 (+593.29)	983.77 (+592.56)
	CIR + Diacylglycerol (C16/C19) + *N*-acyl (C16)	1222.57 (+831.36)	1221.98 (+830.77)	-
LppX	**CSS…EIR**	**2964.46**	-	-
	CSS…EIR + Diacylglycerol (C16/C16)	3515.33 (+550.87)	3514.94 (+550.48)	3514.94 (+550.48)
	CSS…EIR + Diacylglycerol (C16/C19)	3557.42 (+592.96)	-	3556.96 (+592.50)
	CSS…EIR + Diacylglycerol (C16/C19) + Hexose	3719.66 (+755.20)	-	3719.00 (+754.54)
	CSS…EIR + Diacylglycerol (C16/C19) + N-acyl (C16)	3795.82 (+831.36)	3795.21 (+830.75)	-
	CSS…EIR + Diacylglycerol (C16/C19) + 2 Hexoses	3881.90 (+917.44)	-	3881.06 (+916.60)
	CSS…EIR + Diacylglycerol (C16/C19) + *N*-acyl (C16) + Hexose	3958.06 (+993.60)	3957.28 (+992.82)	-
	CSS…EIR + Diacylglycerol (C16/C19) + *N*-acyl (C16) + 2 Hexoses	4120.30 (+1155.84)	4119.45 (+1154.99)	-

Peptides correspond to the N-terminal AspN-digested/tryptic peptides of LprF, LpqH, LpqL and LppX upon cleavage of the signal peptide by LspA. Mass differences to the corresponding unmodified peptide (bold number) due to modifications are given in parentheses. Observed modifications are: diacylglycerol with C16 fatty acid and C16 fatty acid (+550.87 Da). Diacylglycerol with C16 fatty acid and tuberculostearic acid (C19:0) (+592.96 Da), plus one hexose (+162.24 Da, Σ = 755.20 Da) or two hexoses (+324.48 Da, Σ = 917.44). Diacylglycerol with C16 fatty acid and C19:0 fatty acid (+592.96 Da) plus *N*-acyl with C16 fatty acid (+238.40 Da, Σ = 831.36), *N*-acyl with C16 fatty acid plus one hexose (+162.24 Da, Σ = 993.6 Da) or two hexoses (+324.48 Da, Σ = 1155.84 Da). Or diacylglycerol with C16 fatty acid and C19:0 fatty acid (+592.96 Da) plus *N*-acyl with C19:0 fatty acid and hexose (+280.49 Da +162.24 Σ = 1035.69).

Lipoproteins in slow-growing Mycobacteria are N*-acylated with C16 or C19 fatty acids*

Since *N*-acylation was shown to be a common motif in lipoproteins of high GC-rich Gram-positive *M. smegmatis* (Brülle *et al.*, 2010; Tschumi *et al.*, 2009), we proposed Lnt modification also taking place in slow-growing mycobacteria. This proposal was based on the observation that *M. tuberculosis* apolipoprotein *N*-acyltransferase Ppm1 could complement a *M. smegmatis lnt* mutant (Tschumi *et al.*, 2009).

In *M. bovis* BCG, differences in molecular mass of about 831.36 Da for LprF, LpqH, LpqL and LppX, 993.60 Da for LppX, 1035.69 Da for LprF and 1155.84 Da for LppX between the experimentally determined peptide and unmodified N-terminal peptide were found (Table 1). These differences indicated posttranslational modifications of lipoproteins by Lgt, LspA and Lnt. The difference in molecular mass of 831.36 Da points to a modification with diacylglyceryl residue with ester-linked C16 and C19 fatty acid and amide-linked C16 fatty acid. The difference of 993.60 Da indicates a modification with diacylglyceryl residue with ester-linked C16 and C19 fatty acid, amide-linked C16 fatty acid and a glycosylation with one hexose on an *O*-glycosylation site in the N-terminal peptide of LppX. The difference of 1155.84 Da points to a modification with diacylglyceryl residue carrying ester-linked C16 and C19 fatty acid, amide-linked C16 fatty acid and a glycosylation with two hexoses. The difference in molecular mass of 1034.32 Da suggests a modification of LprF with diacylglyceryl residue carrying ester-linked C16 and C19 fatty acid, amide-linked C19 fatty acid and a glycosylation with one hexose (Table 1). Moreover, differences in molecular mass of about 550.87 Da for LppX and 592.96 Da for LpqH, LpqL and LppX were found, both indicating (Lgt and LspA, but not Lnt modified peptides carrying) a diacylglycerol modification with ester-linked C16 and C16 or ester-linked C16 and C19 fatty acid, respectively.

All modifications we estimated from MS data were confirmed by MS/MS (Table 2). Modifications with diacylglyceryl residue were confirmed by eliminations of fragments with the mass of 626.53 Da (C16/C19), corresponding to the elimination of a diacylthioglyceryl carrying C16 and C19 fatty acid. The *O*-linked C16 or C19 fatty acids were confirmed by neutral losses of 256.24 Da and 298.29 Da, corresponding to the elimination of palmitic acid or tuberculostearic acid. Further, neutral losses of 328.24 Da and 370.29 Da correspond to the elimination of C16 or C19 fatty acid α-thioglyceryl ester, respectively. Proposed modification with *N*-linked C16 fatty acid was identified by the neutral loss of 307.26 Da which is

TABLE 2. Comparison of experimentally determined eliminations from N-terminal AspN digested/tryptic peptides of LprF, LpqH, LpqL and LppX in the MALDI-TOF/TOF spectra of parental and mutant strain with theoretically calculated eliminations.

modification	eliminated fragment	calculated mass of eliminated fragment [*Da*]	experimentally determined mass of eliminated fragment [*Da*]											
			parental strain					Δ*lnt*						
			LprF		LpqH	LpqL	LppX	LprF		LpqH			LpqL	LppX
			C16/C19 C16	C16/C19 C19	C16/C19C 16	C16/C19C 16		C16/C16	C16/C19	C16/C16	C16/C18	C16/C19	C16/C19	
O-linked palmitoyl (C16)	Palmitic acid	**256.24**	256.5	-	256.3	256.3	n.d. *	-	-	256.2	256.1	256.3	256.3	n.d. *
O-linked oleyl (C18)	Oleic acid	**282.24**	-	-	-	-	n.d. *	-	-	-	282.4	-	-	n.d. *
O-linked tuberculostearyl (C19)	Tuberculostearic acid	**298.29**	-	-	298.3	298.3	n.d. *	-	-	-	-	298.3	298.4	n.d. *
N-linked palmitoyl (C16) + Didehydroalanine	Palmitamide + Didehydroalanine	**307.26**	-	306.6	-	-	n.d. *	-	-	-	-	-	-	n.d. *
N-linked tuberculostearyl (C19) + Didehydroalanine	Tuberculostearinamide + Didehydroalanine	**349.31**	349.8	-	-	-	n.d. *	-	-	-	-	-	-	n.d. *
Diacylglyceryl (C16/C16)	Diacylhioglyceryl (C16/C16)	**584.44**	-	-	-	-	n.d. *	583.3	-	-	-	-	-	n.d. *
Diacylglyceryl (C16/C18)	Diacylhioglyceryl (C16/C18)	**610.52**	-	-	-	-	n.d. *	-	-	-	-	-	-	n.d. *
Diacylglyceryl (C16/C19)	Diacylhioglyceryl (C16/C19)	**626.53**	625.9	626.7	626.7	626.6	n.d. *	-	626.7	-	-	626.6	626.7	n.d. *
	C16 fatty acid α-thioglyceryl ester	**328.24**	-	-	328.4	328.3	n.d. *	-		-	-	-	-	n.d. *
	C19 fatty acid α-thioglyceryl ester	**370.29**	-	-	370.5	370.3	n.d. *	-	369.8	-	-	-	370.4	n.d. *
Hexose	Hexose	**160.76**	161.62	-	-	-	n.d. *	-	162.9	-	-	-	-	n.d. *

* MALDI-TOF/TOF data for LppX from *M. bovis* BCG were not determined, since MS data of LppX from this study are comparable with data of LppX from *M. smegmatis* (A.Tschumi 2009).

consistent with the elimination of palmitamide plus didehydroalanine. Glycosylations in the tryptic or AspN-digested N-terminal peptides at other amino acids than the conserved cysteine were confirmed by the eliminations of fragments of 162.24 Da for each hexose. (Note, since MS data of LppX from this study are comparable with data from our recent study in *M. smegmatis* (Tschumi *et al.*, 2009), MS/MS data for LppX were not further determined).
Previous structure analyses of lipoprotein modifications in *M. smegmatis* recovered C16 and C19 moieties as ester-linked acyl residues of the diacylglycerol and C16 fatty acid exclusively as substrate for *N*-acylation (Brülle *et al.*, 2010; Tschumi *et al.*, 2009). However, beside the signal at m/z=3326.828, an additional signal at m/z = 3530.562 was found in the MS of LprF (Fig. 1A). The signal at m/z = 3326.828 corresponds to LprF modified with a diacylglyceryl residue carrying ester-linked C16 and C19 fatty acid and *N*-linked C16 fatty acid. Eliminated fragments in MS/MS analysis of the signal m/z = 3530.562 (Fig. 1B) confirmed a modification with diacylglyceryl residue carrying ester-linked C16 and C19 fatty acid, *N*-linked C19 fatty acid and a hexose. The neutral loss of 625.89 Da from the ion at m/z = 3368.508 corresponds to the elimination of diacylthioglyceryl carrying both *O*-linked C16 and C19 fatty acids. In addition, the neutral loss of 349.82 Da from m/z = 2742.615 corresponds to the elimination of tuberculostearinamide plus didehydroalanine. This fragmentation pattern shows that the +1 cysteine is modified at the sulfhydryl group by a diacylglyceryl residue carrying ester-bound C16 fatty acid and C19 fatty acid and an amide-bound C19 fatty acid at the cysteine (Fig. 1C).

Generation of an lnt *deletion mutant in* M. bovis *BCG*

Using *E. coli* Lnt as a query in a BLASTp search on a subset of mycobacteria, we identified three open reading frames annotated as polyprenol-monophosphomannose synthase Ppm1, i.e. Rv2051c in *M. tuberculosis*, BCG_2070c in *M. bovis* BCG Pasteur and MSMEG_3860 in *M. smegmatis*, respectively. In *M. tuberculosis* two additional putative homologous open reading frames, Rv2262c and Rv2261c annotated as hypothetical proteins were found (Fig. 2). Both, MSMEG_3860 as well as the N-terminal part of the two-domain protein encoded by Rv2051c are already identified as functional *N*-acyltransferases in mycobacteria (Tschumi *et al.*, 2009). A further search with *M. tuberculosis* Rv2262c/2261c as a query in a BLASTp search identified BCG_2279c as homologue in *M. bovis* BCG Pasteur, whereas no homologue was found in *M. smegmatis*. We used sequence alignment with the Needleman-Wunsch algorithm (http://www.ebi.ac.uk/Tools/psa/emboss_needle) with default settings to compare both

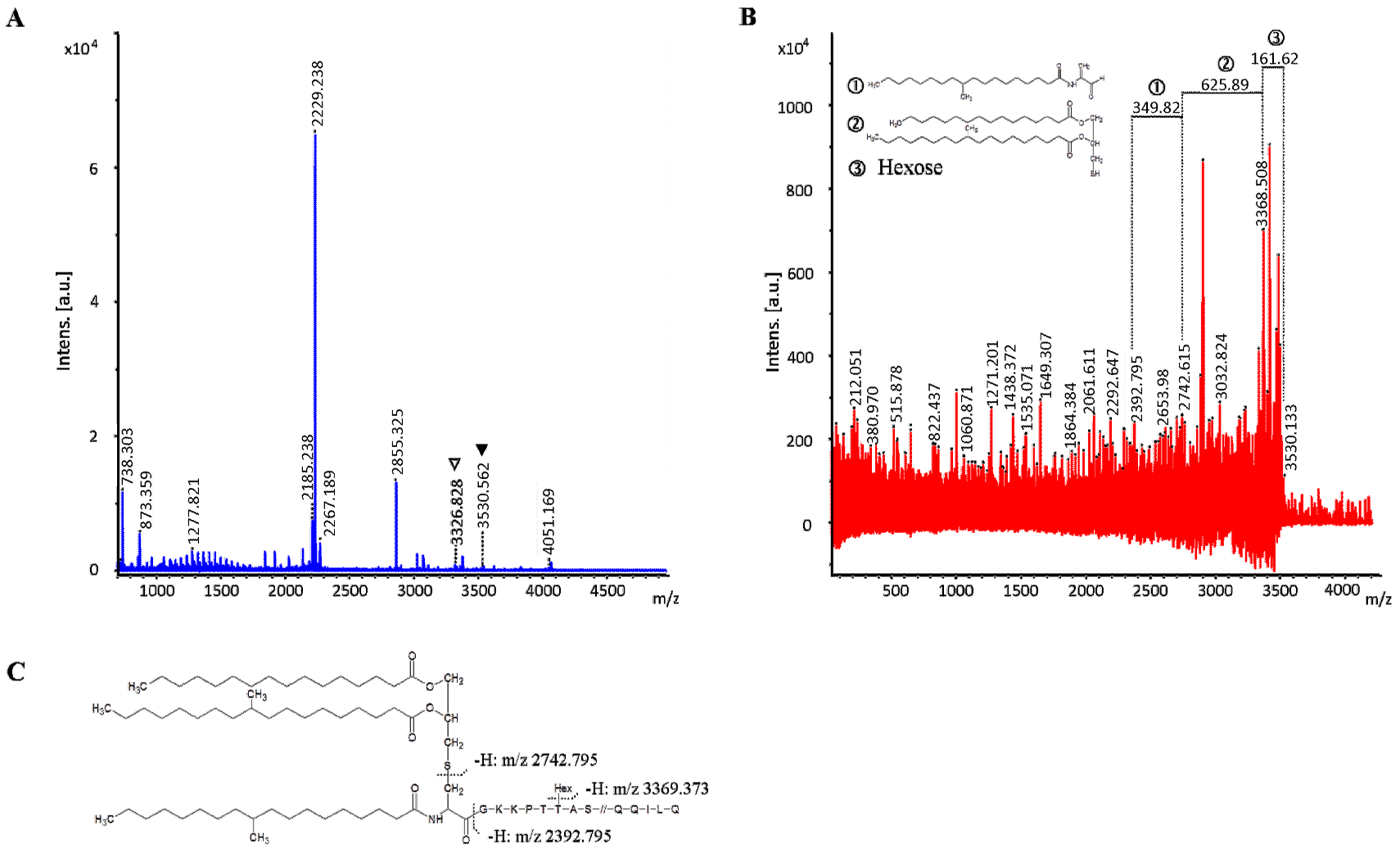


Fig. 1. MALDI-TOF and MALDI-TOF/TOF analysis of the N-terminal peptides of LprF. A. MS analysis of AspN digested peptides of LprF purified from *M. bovis* BCG parental strain. *Filled triangle*, diacylglycerol (C16/C19) + *N*-acyl (C16) modified and glycosylated N-terminal peptide, *open triangle,* diacylglycerol (C16/C19) + *N*-acyl (C19) modified and glycosylated N-terminal peptide B. MS/MS analysis of N-terminal peptides of LprF from *M. bovis* BCG parental strain. Eliminated fragments of LprF modifications are shwon in the upper part of the spectrum. ① Tuberculostearinamide + Didehydroalanine, ② Diacylthioglyceryl (C16/C19), ③ Hexose. C. Schematic drawing of the modified +1 cysteine with the cleavage sites of each identified *m/z* signal.

M. bovis ORFs to *E. coli lnt*, *M. tuberculosis lnt* Rv2051c, as well as *M. tuberculosis* Rv2262c/2261c sequences. Pairwise sequence alignment revealed the highest sequence identity (100%) between BCG_2070c and Rv2051c from *M. tuberculosis*. Interestingly, pairwise sequence alignment of BCG_2279c and Rv2262c/2261c reveals that both sequences differ by a 2bp insertion in Rv2262c (see additional file 3). This leads to a stop codon and initiation of Rv2261c with codon ttg. BCG_2279c does not have this insertion and therefore encodes only one protein. We confirmed this polymorphism by sequencing corresponding regions of *M. tuberculosis* and *M. bovis* BCG genomes. We also used protein sequence alignment with the Needleman-Wunsch algorithm (http://www.ebi.ac.uk/Tools/psa/emboss_needle) and ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with default settings to analyze the conservation of essential residues (see additional file 4).

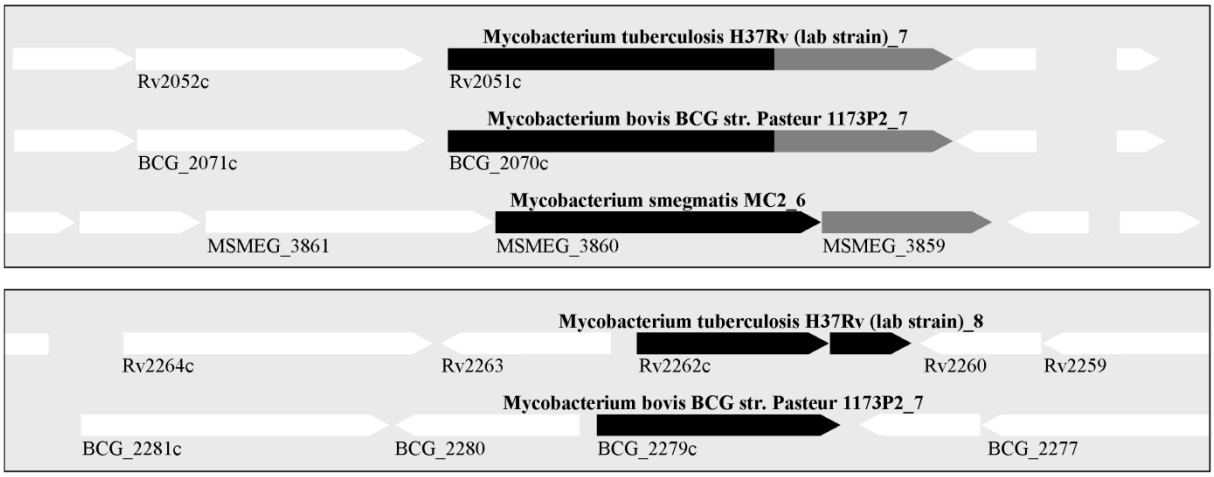


Fig. 2: A comparison of the genomic region of Lnt homologues in mycobacteria. Black bars/arrows indicate Lnt homologues. A second domain is fused to the *lnt* domain in *M. tuberculosis* Rv2051c, and *M. bovis* BCG BCG_2070c (grey arrows) and is homologous to *M. smegmatis* MSMEG_3859 (grey arrow). White arrows indicate orientation of surrounding genes.

BCG_2070c and Rv2051c showed conservation of 14 among 23 residues required for optimal activity of *E. coli* Lnt and conservation of the three essential residues of the catalytic triad of *E. coli* Lnt i.e. E267, K335, C387 (see additional file 1) (Vidal-Ingigliardi *et al.*, 2007). For comparison, the alignment of BCG_2279c and Rv2262c/2261c with *E. coli* Lnt also showed conservation of 13 or 12 (in Rv2262c/2261c *E. coli* P346 is altered from proline to leucine) among the 23 residues of *E. coli* Lnt. However, different residues among the 23 were conserved (see additional file 1). In BCG_2279c and Rv2262c/2261c it revealed that essential residue C387 of the catalytic triad is altered from cysteine to serine. C387 is essential for Lnt-activity and transfer of the acyl residue to the apo-lipoprotein in *E. coli.* However, it has been shown that a Lnt (C387S) mutant also formed an acyl-enzyme intermediate as the wildtype

Lnt C387, but did not have any detectable Lnt activity (Buddelmeijer & Young, 2009; Vidal-Ingigliardi *et al.*, 2007). Moreover, it is noteworthy that the residues of the catalytic triad are separated on two different ORFs encoded by Rv2262c/2261c in *M. tuberculosis*. Beside the three essential residues of the catalytic triad, four other essential residues W237, E343, Y388 and E389 are absolutely required for Lnt function. Among these seven essential residues, five residues are conserved in *M. tuberculosis* Rv2051c, Rv2262c/2261c and *M. bovis* BCG BCG_2070c, BCG_2279c Lnt homologues.

In summary, homology searches and comparison of essential residues in the putative Lnts revealed only small differences and it may be hypothesized that both BCG_2070c and BCG_2279c are functional *N*-acyltransferases. BCG_2070c is identical to an ORF with proven *N*-acyltransferase activity since *M. tuberculosis* Lnt complemented the *M. smegmatis* *lnt* deletion mutant and all three residues of the catalytic triad essential for Lnt function in *E. coli* are conserved. Lnt activity of BCG_2279c may be buried by the Lnt activity of BCG_2070c. Therefore we generated a BCG_2070c *lnt* deletion mutant and characterized lipoprotein modifications in the mutant. The *lnt* deletion mutant was constructed by transformation of *M. bovis* BCG with the suicide plasmid pMCS5-*rpsL-hyg*-Δ*lnt*BCG applying *rpsL* counter-selection strategy, a powerful tool to generate deletion mutants in mycobacteria (Sander *et al.*, 1995; Sander *et al.*, 2001). The mutant strain resulting from allelic exchange is referred to as *M. bovis* BCG Δ*lnt*. Deletion of *lnt* was verified by Southern blot analysis using a 5'*lnt* DNA probe (see additional file 5). The probe hybridized to an 8.1-kbp fragment of the parental strain and to a 3.1-kbp fragment of the Δ*lnt* mutant. Moreover, a complemented mutant strain was constructed by transformation of *M. bovis* BCG Δ*lnt* mutant with complementation vector pMV361-hyg-*lnt*BCG_2070c expressing *M. bovis* BCG BCG_2070c. The complemented strain is referred to as *M. bovis* BCG Δ*lnt*-*lnt*BCG_2070c.

BCG_2070c is a functional N-*acyltransferase in* M. bovis *BCG*

The four expression vectors pMV261-Gm for hexa-histidine/hemagglutinine tagged LprF, LpqH, LpqL or LppX were transformed into *M. bovis* BCG Δ*lnt* mutant. Recombinant lipoproteins expressed in the four strains were analyzed by Western blot. The apparent molecular masses of the detected proteins correspond to the predicted mass of the recombinant apolipoproteins/mature lipoproteins. Eventually the prepro-/pro-lipoprotein forms, whose sizes are increased by 2-3 kDa due to the presence of the signal peptide, are also detected.

In order to characterize BCG_2070c and eventually residual *N*-aclytransferase activity in *M. bovis BCG*, lipoprotein modifications of LprF, LpqH, LpqL and LppX from Δ*lnt* mutant were analyzed at the molecular level. In Δ*lnt*, signals with molecular masses indicating Lgt- and LspA- modified and glycosylated peptides were found. The differences in molecular mass of 550.87 Da for LprF, LpqH and LppX and 576.91 Da for LprF and LpqH between the experimentally found peptide and the unmodified N-terminal peptide (Table 1) indicate (Lgt and LspA, but not Lnt modified peptides carrying) a diacylglycerol modification carrying ester-linked C16 and C16 or ester-linked C16 and C18 fatty acid, respectively. The differences in molecular mass of 592.96 Da for LprF, LpqH, LpqL and LppX refer to a diacylglycerol modification with ester-linked C16 and C19 fatty acid. The differences in molecular mass of 755.20 Da for LprF and LppX refer to a diacylglycerol modification with ester-linked C16 and C19 fatty acid plus glycosylation with one hexose (592.96 Da + 162.24 Da). The difference in molecular mass of 917.90 Da for LppX refers to a diacylglycerol modification with ester-linked C16 and C19 fatty acid plus modification with two hexoses (592.96 Da + 162.24 Da + 162.24 Da).

In contrast to the MS from parental strain, no molecular masses which we calculated for modifications with three fatty acids were found in the Δ*lnt* mutant strain. In particular, the differences in molecular mass of 238.4 Da (831.36 Da - 592.96 Da) or 280.49 Da (1035.69 Da - 162.24 Da - 592.96 Da) between the C16/C19/C16 or C16/C19/C19 triacylated modification found in the parental strain and the corresponding estimated C16/C19 modification in the Δ*lnt* mutant indicate a lack of *N*-acylation with a C16 or C19 fatty acid in the Δ*lnt* mutant. In MS/MS analysis, this indication of missing *N*-acylation in the mutant was confirmed by identification of the estimated modifications and information about its linkage (Table 2). Modifications with C16/C19 diacylglyceryl residue were confirmed by eliminations of fragments with the molecular mass of 626.53 Da, corresponding to the elimination of a diacylthioglyceryl carrying C16 and C19 fatty acid. The *O*-linked C16 or C19 fatty acids were confirmed by neutral losses of 256.24 Da or 298.29 Da, corresponding to the elimination of palmitic acid or tuberculostearic acid, respectively. Further, the neutral loss of 370.29 Da corresponds to the elimination of C19 fatty acid α-thioglyceryl ester. A glycosylation at other amino acids than the conserved cysteine was confirmed by the release of a fragment of 162.24 Da for a hexose. These findings indicate that *N*-acylation is not a prerequisite for glycosylation. As mentioned before, only diacylglyceryl residues composed of a C16 and a C19 fatty acid were identified in mycobacterial lipid anchors so far (Brülle *et al.*, 2010; Tschumi *et al.*, 2009). However, the eliminations of fragments with the molecular mass of

584.44 Da or 256.24 Da, corresponding to the elimination of diacylthioglyceryl and palmitic acid, give evidence for modifications with diacylglyceryl residue carrying C16 and C16 fatty acids. Moreover, estimated diacylglycerol modifications carrying C16 and C18 fatty acids were confirmed by neutral losses of fragments with the molecular mass of 256.24 Da and 282.44 Da, corresponding to the elimination of palmitic and oleic acid. In complemented mutant Δ*lnt-lnt*BCG_2070c, lipoproteins LprF and LppX were triacylated and glycosylated (see additional file 6 and additional file 7). This confirmed that BCG_2070c restored the BCG_2070c mutant.

The absence of *N*-acylation of the four analyzed lipoproteins in the Δ*lnt* mutant and the complementation of the mutant provide strong evidence that BCG_2070c is the only functional apolipoprotein *N*-acyltransferase that modifies these lipoproteins with an amide-linked fatty acid in *M. bovis* BCG. In addition, it demonstrates that BCG_2279c is not able to adopt or substitute *N*-acylation of the four lipoproteins in the Δ*lnt* mutant.

DISCUSSION

Lipoproteins are present in all bacterial species, but their biogenesis and lipid moieties differ, especially between Gram-negative and Gram-positive bacteria. The three enzymes involved in lipoprotein biosynthesis, namely Lgt, LspA and Lnt first were identified in *E. coli*. Therefore, the lipoprotein biosynthesis pathway in *E. coli* is intensively studied and well described (Okuda & Tokuda, 2011). Mycobacteria are classified as Gram-positive bacteria, but their lipoprotein biosynthesis pathway resembles that of Gram-negative bacteria. The discovery of Lnt in mycobacteria and the identification of lipoprotein *N*-acylation in *M. smegmatis* renewed interest within the field of mycobacterial lipoprotein research. The evidence of triacylated lipoproteins in mycobacteria refuted the long held assumption, that *N*-acylation is restricted to Gram-negative bacteria. Thus, the acylation with three fatty acids is a common feature of mycobacterial and *E. coli* lipoproteins. But, mycobacterial lipoproteins differ from *E. coli* lipoproteins with respect to the fatty acids used for the triacylation. Mycobacteria-specific fatty acid 10-methyl octadecanoic acid (tuberculostearic acid) is uniquely found in lipoproteins of *M. smegmatis* (Brülle *et al.*, 2010; Tschumi *et al.*, 2009).

All three enzymes of the lipoprotein biosynthesis pathway, Lgt, LspA and Lnt are essential in Gram-negative, but not in Gram-positive bacteria. However, in *M. tuberculosis, lgt*, the first enzyme of the lipoprotein biosynthesis pathway is essential. A targeted deletion of *lgt* was not

possible (Tschumi *et al.*, 2012). In contrast, an *lspA* deletion mutant was viable, but the mutant strain showed a reduced number of CFU in an animal model and induced hardly any lung pathology. This confirmed a role of the lipoprotein biosynthesis pathway in pathogenesis of *M. tuberculosis* (Rampini *et al.*, 2008; Sander *et al.*, 2004).

Lipoproteins itself are well known virulence factors in pathogenic bacteria. *M. tuberculosis* lipoproteins in particular have been shown to suppress innate immune responses by TLR2 agonist activity (Drage *et al.*, 2009). Newest data also show that lipoproteins constitute the main proteinaceous content of membrane vesicles released by pathogenic mycobacteria and that they are highly immunogenic (Ziegenbalg *et al.*, 2013). Several *M. tuberculosis* mutants deficient in individual lipoproteins are attenuated in virulence as shown for LppX (Camacho *et al.*, 1999), LprG (Bigi *et al.*, 2004) and LpqH (Henao-Tamayo *et al.*, 2007). Recently, a *M. tuberculosis* deletion mutant, defective in lipoprotein LpqS showed attenuation in macrophages (Sakthi & Narayanan, 2013). Despite the important role of *M. tuberculosis* lipoproteins in immunogenicity and pathogenicity and all the achievements in knowledge about the lipoprotein modification in apathogenic *M. smegmatis*, still little is known about the molecular structure of lipoproteins in pathogenic mycobacteria. The elucidation of lipoprotein structure can build the fundamental knowledge for future development of lipoprotein based subunit vaccines and antitubercular drugs targeting enzymes of the lipoprotein synthesis pathway (Gowthaman *et al.*, 2012). Therefore we extended our research in lipoprotein modifications to slow-growing mycobacteria. Most of the pathogenic mycobacteria and the tuberculosis vaccine strain *M. bovis* BCG belong to this sub-group.

In the present study, we investigated the lipid moieties of four mycobacterial lipoproteins representing lipoproteins with different functions. By MALDI-TOF/TOF analyses of a Trypsin digest of purified LpqH, LpqL and LppX and an AspN digest of purified LprF, we unambiguously identified modifications at the universally conserved cysteine in the parental strain. All four proteins were found to be triacylated carrying a thioether-linked diacylglyceryl residue with C16 and C19 fatty acid (C16/C19) to the sulfhydryl group of the lipobox cysteine and an amide-linked C16 fatty acid. Whether the fatty acids of the diacylglyceryl residue are in the S_n1 or S_n2 position could not be determined by mass spectrometry and therefore currently remains elusive. In LprF, a novel triacylation with C16/C19 diacylglycerol and C19 *N*-acyl was identified. This differs from previous lipoprotein analyses in *M. smegmatis*, where C16 fatty acid was the single substrate for Lnt (Brülle *et al.*, 2010; Tschumi *et al.*, 2009). Likewise, it shows that mycobacteria not only use mycobacteria-specific fatty acids for diacylglycerol modification, but also for *N*-acylation. Lipoprotein modifications with acyl

residues of different length, stiffness and bulkiness may influence membrane fluidity and localization of lipoproteins. In *Francisella novicida*, an environmentally regulated membrane remodelling directed by multiple alleles of the lipid A-modifying *N*-acyltransferase enzyme is reported. By incorporation of shorter or longer *N*-acyl fatty acid chains to the outer membrane lipid A, the bacterium regulates the maintenance of membrane fluidity and integrity (Li *et al.*, 2012). Therefore, it is obvious to speculate a similar important role of the C19 *N*-acyl lipoprotein modification for mycobacteria in terms of adaptations to environmental alterations or specific bacterial conditions. Adaptation of lipoprotein modification to environmental conditions has been shown for *S. aureus*. A combination of conditions including acidic pH and post-logarithmic growth phase induced the accumulation of diacylated lipoproteins (Kurokawa *et al.*, 2012a).

By the usage of C19 fatty acid, mycobacterial Lnt strongly differs in substrate specificity from *E. coli* Lnt. *E. coli* Lnt utilizes all three major phospholipids of *E. coli* phosphatidylethanolamine, phosphatidylglycerol and cardiolipin as its fatty acid source *in vivo* (Gupta *et al.*, 1991). Subsequent analysis revealed that both the phospholipid head group and its acyl chain composition affect *N*-acyltransferase activity *in vitro* (Hillmann *et al.*, 2011). *E. coli* Lnt incorporates palmitic (C16) fatty acids from the S_n1 position of phospholipids to diacylated lipoproteins (Jackowski & Rock, 1986). In mycobacterial phospholipids the S_n1 position is esterified principally with octadecanoic or tuberculostearic acid (C18 related fatty acids), whereas palmitic acid (C16) is mainly located at the S_n2 position (Okuyama *et al.*, 1967). Based on this and the fact, that palmitic acids were used for *N*-acylation of lipoproteins in *M. smegmatis* (Brülle *et al.*, 2010; Tschumi *et al.*, 2009), Nakayama *et al.* proposed that *M. smegmatis* Lnt uses fatty acids from the S_n2 position as substrates and therefore has a different specificity than *E. coli* Lnt (Nakayama *et al.*, 2012). This specificity obviously is different in *M. bovis* BCG. Our results provide strong evidence, that not only palmitic acid from the S_n2 position, but also tuberculostearic acid (C19), a fatty acid from the S_n1 position of phospholipids is transferred by Lnt (Okuyama *et al.*, 1967).

Lipoproteins are recognized by TLR2 in association with TLR1 or TLR6. While diacylated lipoproteins carrying the *S*-diacylglyceryl residue are recognized by TLR2/6 heterodimers, triacylated lipoproteins carrying the additional *N*-acyl are recognized by TLR1/2 heterodimers. The two ester-bound fatty acids are inserted into a pocket in TLR2 while the amide-bound fatty acid is inserted into a hydrophobic channel in TLR1. Therefore the *N*-acyl of the lipoprotein is indispensable for the heterodimerization of TLR2 and TLR1 and thus the initiation of TLR2/1 signaling (Jin *et al.*, 2007; Kang *et al.*, 2009). Recent investigations

indicate that TLR1 polymorphisms are associated with resistance towards bacterial pathogens, including *M. tuberculosis* (Azad *et al.*, 2012; Mayerle *et al.*, 2013). It may be hypothesized that the modification of lipoproteins with particular fatty acids plays a crucial role for lipoprotein function, its retention in a membrane, and interaction with TLRs. However, whether the *N*-acylation with C19 fatty acid is only characteristic for LprF or also for other lipoproteins and whether it is a feature of *M. bovis* BCG Lnt remains to be investigated.

Beside the triacylated forms, also diacylated forms of the N-terminal peptide were found in proteins from the parental BCG strain. A modification with C16/C19 diacylglycerol was found in LpqL and a C16/C16 diacylglycerol was found in LppX. These molecules probably indicate N-terminal peptides from unmature proteins which have not converted to mature lipoproteins by Lnt yet.

Lipoproteins from *M. bovis* BCG were identified to be triacylated at their N-terminus which suggests that *N*-acylation by an Lnt enzyme also exists in slow-growing mycobacteria. We identified the open reading frame, encoding the Lnt enzyme responsible for the *N*-acylation. *M. bovis* BCG Pasteur genome analysis revealed two open reading frames BCG_2070c and BCG_2279c homologous to *E. coli* Lnt. Our biochemical analyses of four lipoproteins expressed in a BCG_2070c Δ*lnt* mutant demonstrated that BCG_2070c is the major if not the only functional mycobacterial Lnt in *M. bovis* BCG. When we subjected lipoproteins LprF, LpqH, LpqL and LppX expressed in the Δ*lnt* mutant to MALDI-TOF/TOF analyses, none of the proteins was found to be *N*-acylated. All four proteins were found to be only diacylated in contrast to the triacylated proteins in the parental strain. Diacylglyceryl residues composed of C16/C19 fatty acid, C16/C16 fatty acid or C16/C18 were found. Hereby the usage of oleic acid as a substrate for lipoprotein modification in mycobacteria, to our knowledge is shown for the first time.

We showed that the lack of BCG_2070c results in a failure of lipoprotein *N*-acylation and that BCG_2279c is not able to compensate Lnt function. BCG_2279c has a C to S amino acid substitution in C387, a residue essential for Lnt function in *E. coli*. In *E. coli*, a C387 alteration absolutely abolishes Lnt function, because this residue is part of the catalytic triad of Lnt (Vidal-Ingigliardi *et al.*, 2007). Alterations in BCG_2279c therefore could account for its inactivity as Lnt. But we cannot exclude that BCG_2279c is a second Lnt particularly active under specific growth conditions. Alternatively, BCG_2279c may act only on a small subset of dozens of putative mycobacterial lipoproteins not yet characterized by MALDI-TOF/TOF.

Streptomyces spp., bacteria closely related to mycobacteria, also encode two Lnt homologues. Deleting *Streptomyces scabies lnt1* and *lnt2* genes individually or in combination revealed that Lnt1 is a functional Lnt sufficient and required for *N*-acylation. Lnt2 could not compensate for the Lnt1 deletion. However, both Lnts seem to be required for efficient lipoprotein *N*-acylation as the lack of Lnt2 alone resulted in a marginal *N*-acylation activity. This implies a subsidiary but inessential role for Lnt2, not directly involved in *N*-acylation of lipoproteins (Widdick *et al.*, 2011). Likewise, an interplay can count for the two Lnt homologues in *M. bovis* BCG. But, in contrast to the Lnts in *S. scabies*, BCG_2279c is missing one of the three essential residues required for Lnt activity in *E. coli.* This, in our opinion diminishes the possibility for BCG_2279c to be an Lnt with *N*-acylation activity and favours a contributive role for it. *In vitro* biochemical assays (Hillmann *et al.*, 2011) with purified BCG_2279c or analyses of a BCG_2279c mutant alone or in combination with BCG_2070c would be required to elucidate this.

Beside the fatty acid modifications, we also identified hexose glycosylations in LprF and LppX. Modifications with one or more glycosyl residues have been shown for several mycobacterial lipoproteins (Brülle *et al.*, 2010; Herrmann *et al.*, 1996; Sartain & Belisle, 2009). *O*-glycosylation occurs at Ser and Thr residues respectively. Although glycosylations of the tryptic or AspN-digested N-terminal peptides of LprF and LppX were identified, the exact glycosylation site within the peptide could not be determined. No glycosylations were found for N-terminal fragments of LpqH and LpqL. This possibly is due to the use of proteases which have cleavage sites close to the N-terminus and therefore the peptide fragment may be too short to include *O*-glycosylation sites. The information about the exact molecular nature and function of the glycosylation is scarce, but its influence on subcellular lipoprotein localization and its protection from proteolytic degradation are proposed (Herrmann *et al.*, 1996; Sartain & Belisle, 2009). In *B. subtilis* lipoprotein glycosylation is discussed to control a lipoprotein "shaving" mechanism and thus their release into the culture medium (Tjalsma & van Dijl, 2005). In our study, glycosylations were found also in lipoproteins from the Δ*lnt* mutant, demonstrating that *N*-acylation is not a prerequisite for glycosylation. Lnt independent glycosylation was also demonstrated in *C. glutamicum* (Mohiman *et al.*, 2012). In *C. glutamicum* Cg-Ppm1 is responsible for glycosylation. Cg-ppm1 (Ppm synthase) and Cg-ppm2 (Lnt) are similar organized as MSMEG_3859 (Ppm synthase) and MSMEG_3860 (Lnt) in *M. smegmatis* (Fig. 2). Deletion of the Lnt domain of BCG_2070c obviously did not abolish Ppm activity encoded in the same ORF. Of note, Lnt is dispensable while Ppm is essential in *M. tuberculosis* (Zhang *et al.*, 2012).

In Gram-negative bacteria, the efficient lipoprotein transport to the outer membrane depends on the localization of lipoproteins (Lol) transport system and there is good evidence that *N*-acylation by Lnt facilitates lipoprotein translocation in *E. coli* (Okuda & Tokuda, 2011; Robichon *et al.*, 2005). Lnt is essential in *E. coli*, however deletion of *lnt* was possible upon overexpression of proteins from the Lol system, indicating an important role of *N*-acylation in targeting lipoproteins to the outer membrane (Narita & Tokuda, 2011). Mycobacteria have an outer membrane mycolic acid bilayer (Niederweis *et al.*, 2010; Sutcliffe, 2010; Zuber *et al.*, 2008) and are known to localize lipoproteins to the cell surface (Niederweis *et al.*, 2010). Nevertheless, no mechanisms for translocation or transport systems are identified and whether *N*-acylation and glycosylation, alone or in combination are involved in the translocation of specific lipoproteins to the mycolate layer is not known so far.

In the present study we show that lipoproteins from *M. bovis* BCG, the live vaccine for tuberculosis are triacylated and we identified the lipid modifications at the molecular level. BCG_2070c is a functional homologue of *E. coli* Lnt, but differs in substrate specificity. The identification of *N*-linked tuberculostearic acid shows for the first time, to our knowledge, that mycobacteria-specific fatty acids are used by mycobacterial Lnts.

The antituberculosis drug pipeline is not sufficiently filled and the vaccines used at present do not provide effective protection against tuberculosis in adults. For lipoproteins and their biosynthesis pathway potential implications in *M. tuberculosis* pathogenesis and immunogenicity have been shown. Our results about lipoprotein structure therefore may contribute to provide the knowledge which is required to develop novel vaccines and antituberculosis drugs to eliminate this worldwide epidemic.

CONCLUSIONS

Lipoproteins are triacylated in slow-growing mycobacteria. By MALDI-TOF/TOF analyses lipoprotein modifications in *M. bovis* BCG wildtype and BCG_2070c *lnt* deletion mutant were analyzed at the molecular level. *N*-acylation of lipoproteins was only found in the wildtype strain, but not in the mutant strain, which confirmed BCG_2070c as functional *lnt* in *M. bovis* BCG. Moreover, we identified mycobacteria-specific tuberculostearic acid as further substrate for *N*-acylation in slow-growing mycobacteria.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

JKB designed the study, performed experimental work and drafted the manuscript. AT carried out the genetic engineering of *M. bovis* BCG mutant strain and participated in the MS/MS data analyses. PS conceived of the study, participated in its coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the University of Zurich, Swiss National Foundation (31003A_13705), European Union (EU-FP7 New TBVac No 241745) and Stiftung wissenschaftliche Forschung (SWF). We thank Yolanda Joho-Auchli from the Functional Genomics Center Zurich for MALDI-TOF/TOF analysis and Nienke Buddelmeijer for helpful discussions.

ADDENDUM

Personal contribution to chapter 3

My contribution as first author to this manuscript was as follows:

- Design of the study
- Bioinformatic analysis
- Generation of expression vectors
- Lipoprotein expression controls
- Lipoprotein purification
- MALDI-TOF data analysis
- MALDI-TOF/TOF data analysis
- Writing the manuscript

SUPPLEMENTAL MATERIAL

Supplemental Table S1

TABLE S1. Conservation of essential residues in Lnt homologues.

Escherichia coli **Lnt essential residue***	*Mycobacterium tuberculosis* **Rv2051c**	*Mycobacterium bovis* **BCG BCG_2070c**	*Mycobacterium tuberculosis* **Rv2262c/2261c****	*Mycobacterium bovis* **BCG BCG_2279c**
W74	F	F	F	F
F146	G	G	G	G
W148	+	+	+	+
Q228	+	+	+	+
Q233	+	+	G	G
W237	V	V	D	D
Y249	A	A	T	T
P266	+	+	G	G
E267	**+**	**+**	**+**	**+**
N314	T	T	K	K
Y333	H	H	+	+
K335	**+**	**+**	**+**	**+**
E343	+	+	+	+
P346	+	+	L	**+**
Q372	T	T	A	P
C387	**+**	**+**	**S**	**S**
Y388	W	W	+	+
E389	+	+	+	+
Q424	+	+	+	+
R432	+	+	+	+
A433	+	+	+	+
E435	+	+	+	+
L436	H	H	A	A

Essential residues of the catalytic triad are written in bold letters. Other essential residues absolutely required for function are written in italic letters.

*Vidal-Ingigliardi *et al.* 2007

**Residues above the bar are encoded by Rv2262c, residues below the bar are encoded by Rv2261c.

Supplemental Figure S1

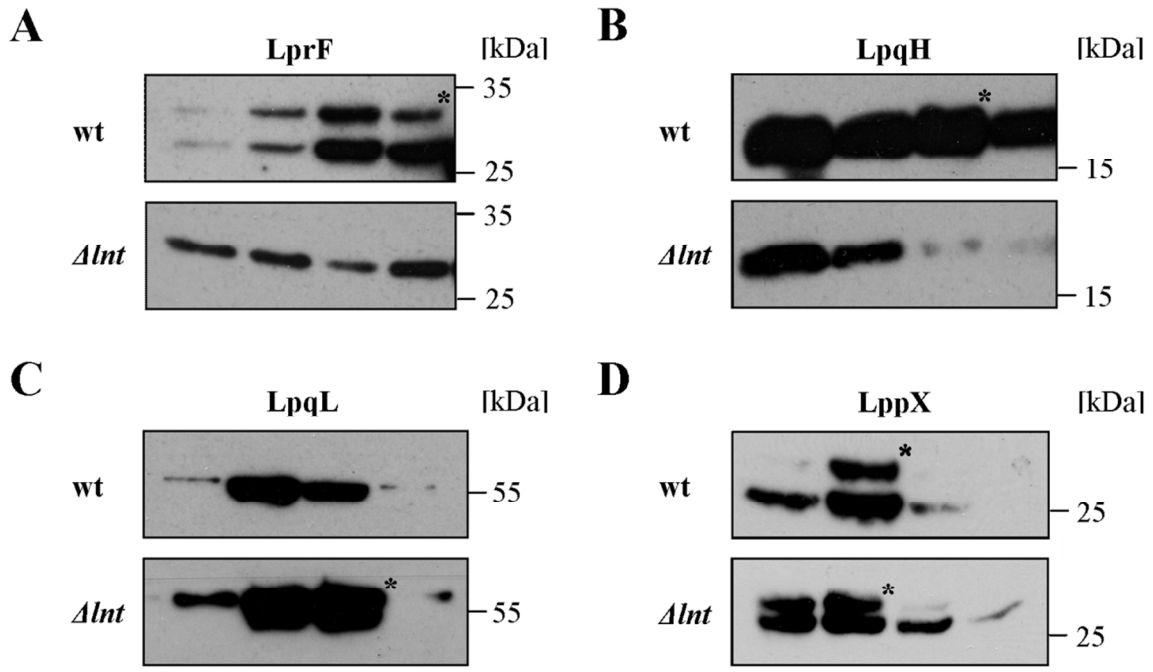


Fig. S1. Western blot analysis of purified lipoproteins of *M. bovis* BCG wildtype and Δ*lnt* mutant strain. Elution fractions from FPLC or HA affinity chromatography were analyzed by Western blot using anti-HA antibody and corresponding secondary antibody conjugated with horseradish peroxidase. A. LprF B. LpqH C. LpqL D. LppX. Bands with a higher molecular weight (labeled with an asterisk) correspond to pre-pro-lipoprotein or pro-lipoprotein, lower bands correspond to apolipoprotein and/or mature lipoprotein, respectively.

Supplemental Figure S2

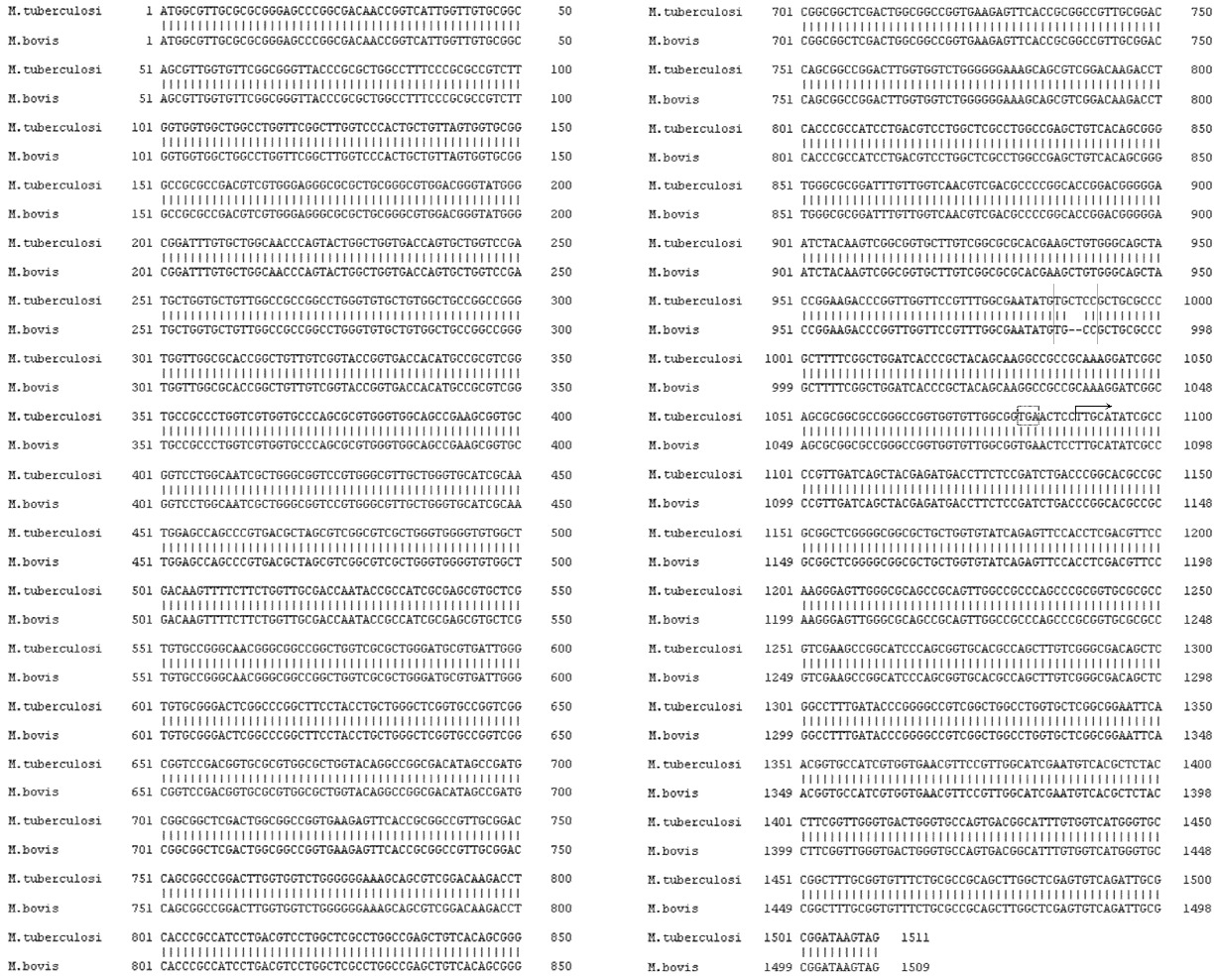


Fig. S2. Sequence alignment of *M. tuberculosis* Rv2262c/Rv2261c and *M. bovis* BCG_2070c using EMBOSS Needle. 2bp difference in sequence is marked with a box. Resulting stop codon TAG of Rv2262c is marked with a box (dotted line) and start of Rv2261c is indicated with an arrow.

Supplemental Figure S3

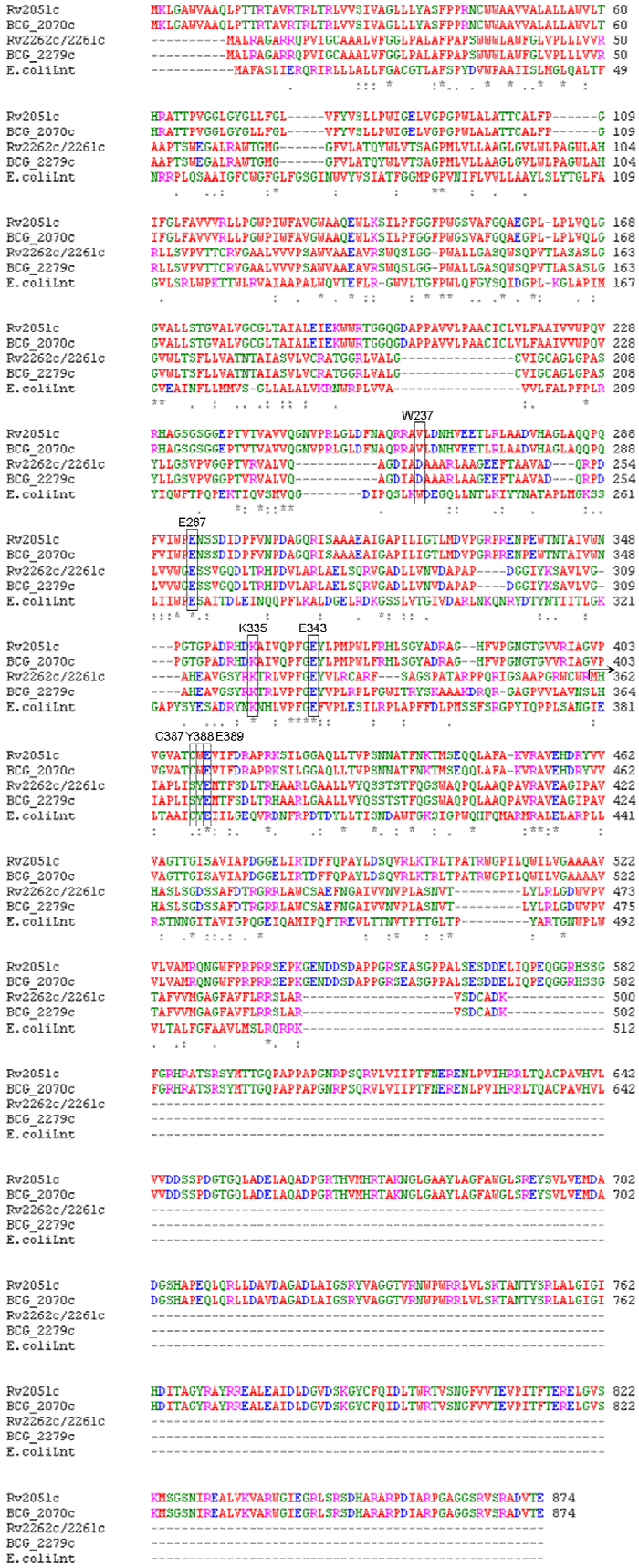


Fig. S3. Multiple sequence alignment of Lnt homologues using Clustal W2. Essential residues required for *E. coli* Lnt function and corresponding residues in the Lnt homologues are marked with a box. Start of Rv2261c sequence is indicated with an arrow.

Supplemental Figure S4

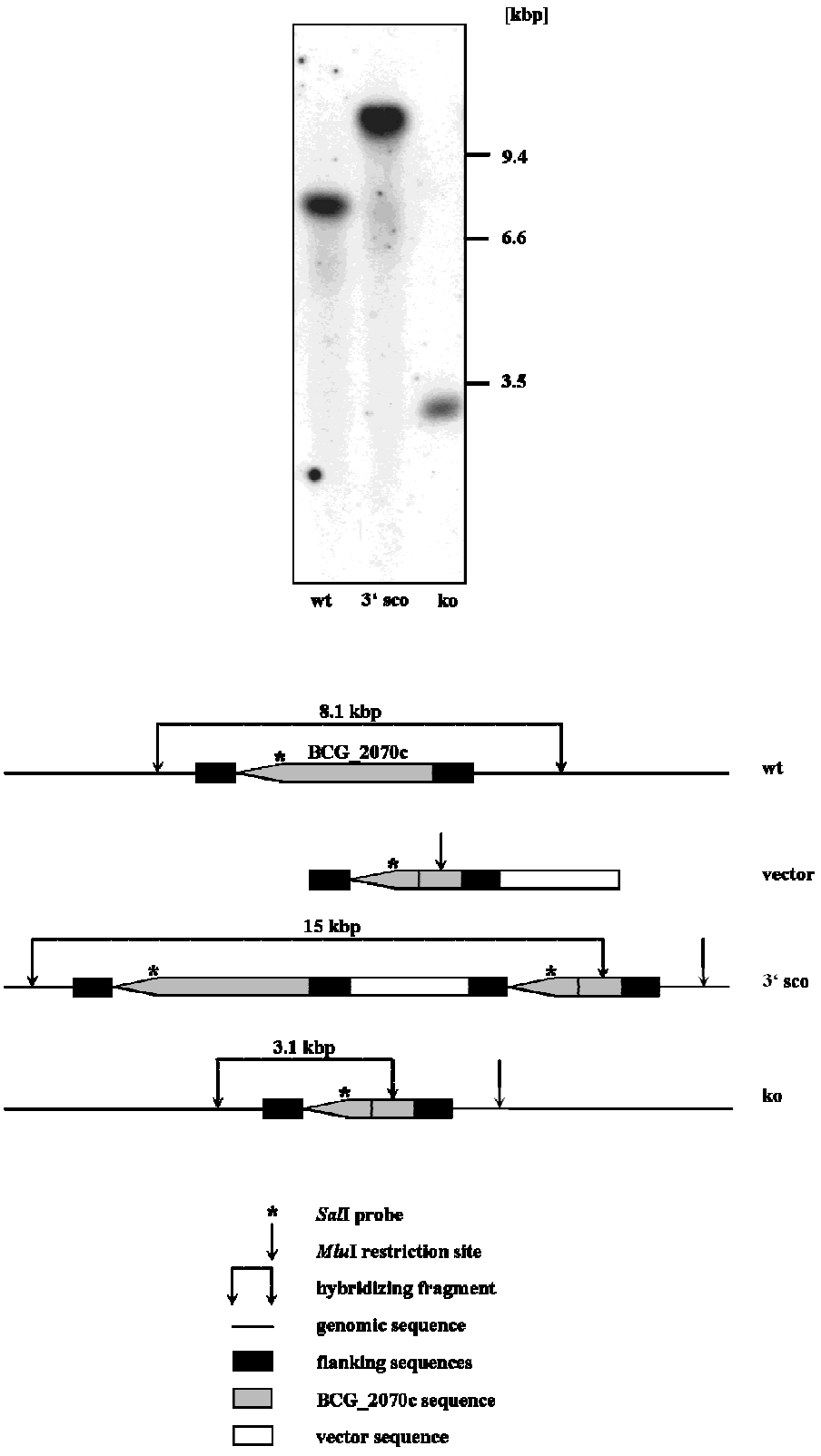


Fig. S4. Disruption of *Mycobacterium bovis* BCG *lnt* (BCG_2070c). Genomic DNAs from *M. bovis* BCG (lane 1), *lnt* single-crossover (3' sco) mutant (lane 2) and Δ*lnt* mutant (lane 3) were digested with *Mlu*I and probed with a 222 bps *Sal*I *lnt* gene fragment. The presence of the 3.1 kbp fragment in the Δ*lnt* knockout compared to the 8.1 kbp fragment in the parental strain demonstrates inactivation of *lnt*.

Supplemental Figure S5

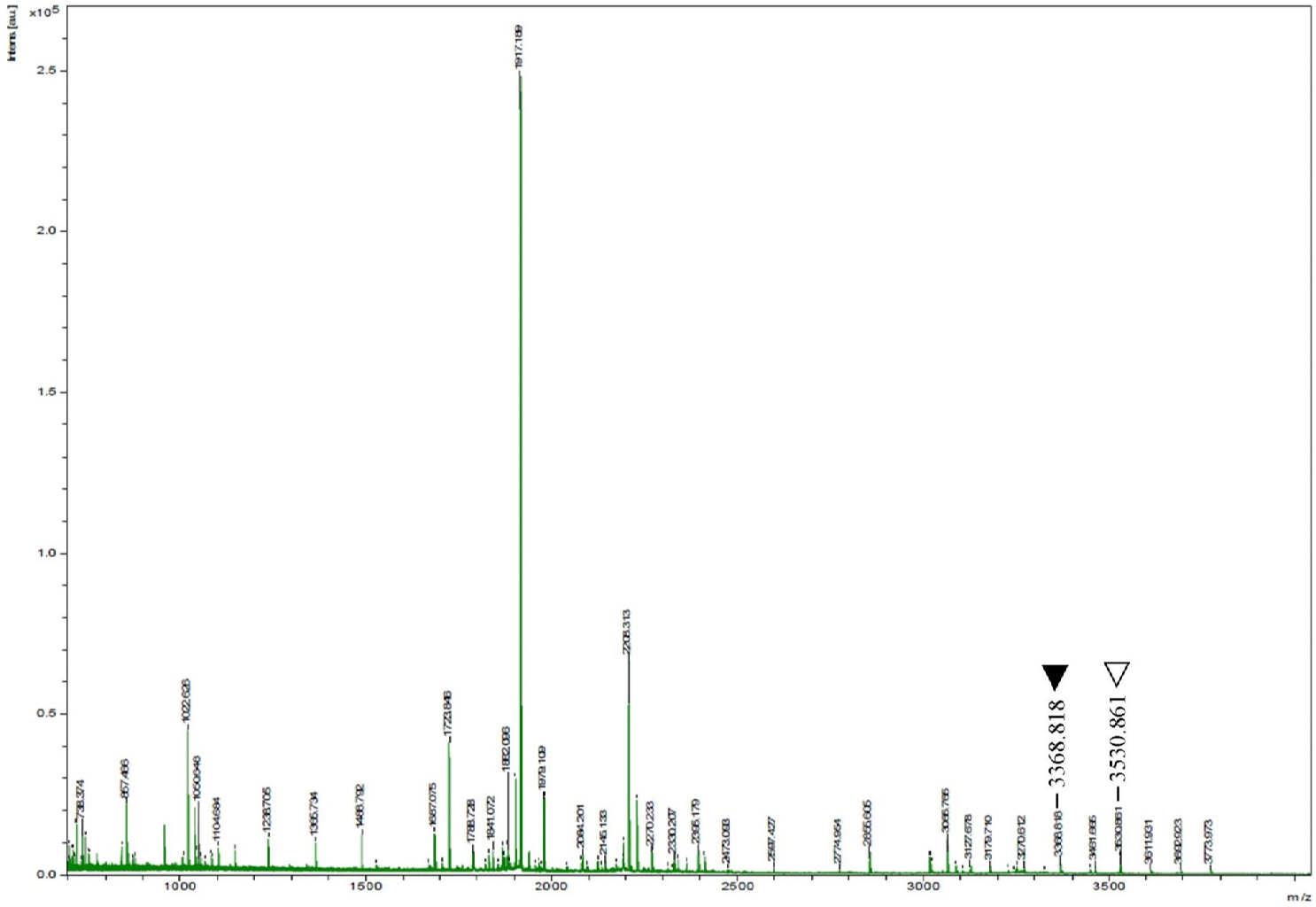


Fig. S5. MALDI-TOF analysis of the N-terminal peptides of LprF. MS analysis of AspN-digested peptides of LprF purified from complemented mutant *Δlnt-lnt*BCG_2070c. *Filled triangle*, diacylglycerol (C16/C19) + *N*-acyl (C19) modified N-terminal peptide, *open triangle*, diacylglycerol (C16/C19) + *N*-acyl (C19) modified and glycosylated N-terminal peptide.

Supplemental Figure S6

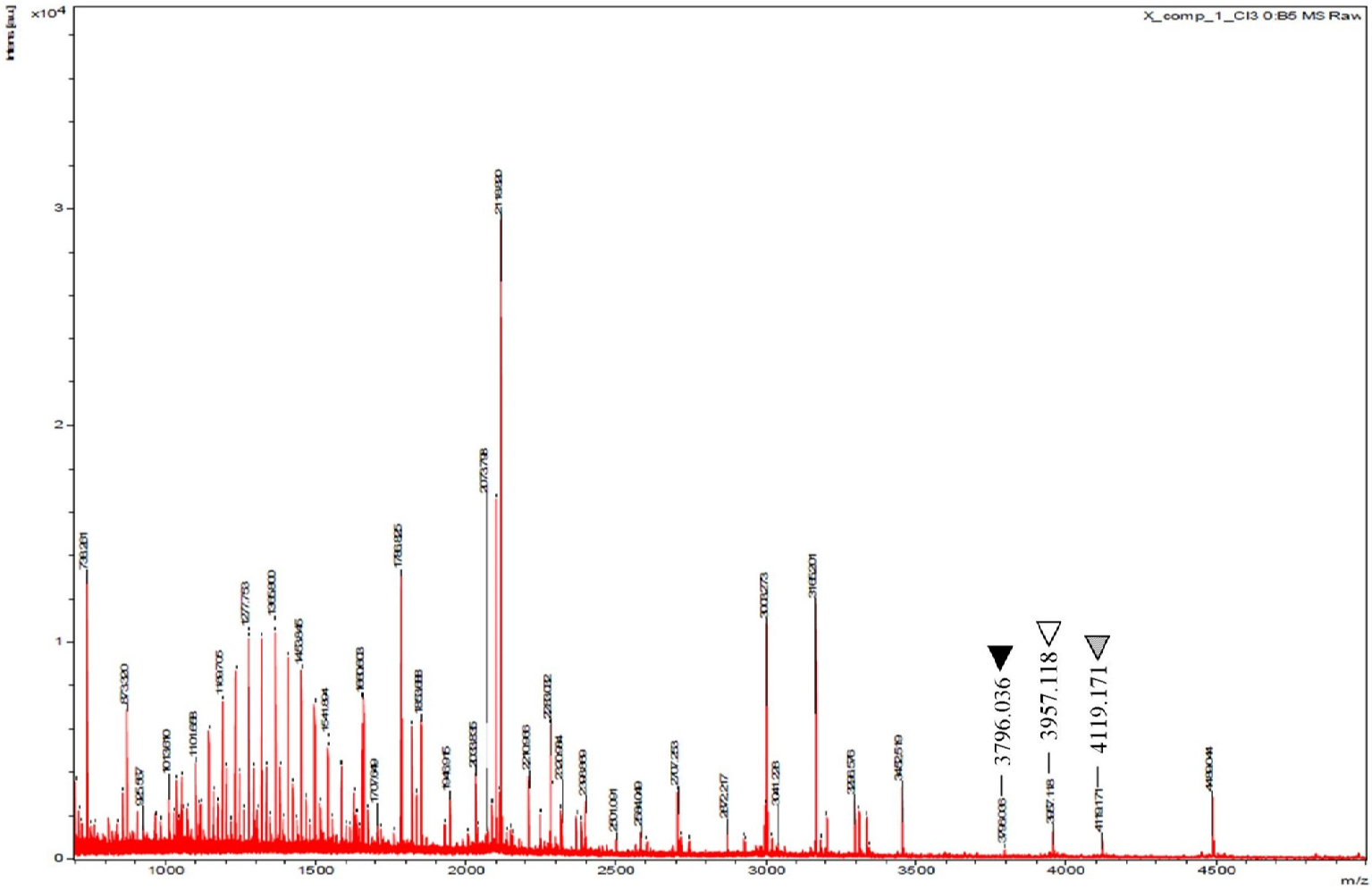


Fig. S6. MALDI-TOF analysis of the N-terminal peptides of LppX. MS analysis of Trypsin-digested peptides of LppX purified from complemented mutant *Δlnt-lnt*BCG_2070c. *Filled triangle*, diacylglycerol (C16/C19) + *N*-acyl (C16) modified N-terminal peptide, *open triangle*, diacylglycerol (C16/C19) + *N*-acyl (C16) modified and glycosylated N-terminal peptide, *shaded triangle*, diacylglycerol (C16/C19) + *N*-acyl (C16) modified and with two hexoses glycosylated N-terminal peptide.

CHAPTER 4

DISSECTING THE COMPLETE LIPOPROTEIN BIOGENESIS PATHWAY IN *STREPTOMYCES SCABIES*

David A. Widdick, Matthew G. Hicks, Benjamin J. Thompson, Andreas Tschumi, Govind Chandra, Iain C Sutcliffe, Juliane K. Brülle, Peter Sander, Tracy Palmer and Matthew I. Hutchings

Molecular Microbiology (2011) 80 (5), 1395-1412

ABSTRACT

Following translocation, bacterial lipoproteins are lipidated by lipoprotein diacylglycerol transferase (Lgt) and cleaved of their signal peptides by lipoprotein signal peptidase (Lsp). In Gram-negative bacteria and mycobacteria, lipoproteins are further lipidated by lipoprotein *N*-acyl transferase (Lnt), to give triacylated lipoproteins. *Streptomyces* are unusual amongst Gram-positive bacteria because they export large numbers of lipoproteins via the twin arginine protein transport (Tat) pathway. Furthermore, some *Streptomyces* species encode two Lgt homologues and all *Streptomyces* species encode two homologues of Lnt. Here we characterise lipoprotein biogenesis in the plant pathogen *Streptomyces scabies* and report that *lgt* and *lsp* mutants are defective in growth and development while only moderately affected in virulence. Lipoproteins are lost from the membrane in an *S. scabies lgt* mutant but restored by expression of *S. coelicolor lgt1* or *lgt2* confirming that both encode functional Lgt enzymes. Furthermore, lipoproteins are *N*-acylated in *Streptomyces* with efficient *N*-acylation dependent on Lnt1 and Lnt2. However, deletion of *lnt1* and *lnt2* has no effect on growth, development or virulence. We thus present a detailed study of lipoprotein biogenesis in *Streptomyces*, the first study of Lnt function in a monoderm bacterium and the first study of bacterial lipoproteins as virulence factors in a plant pathogen.

INTRODUCTION

Bacterial lipoproteins are a specialised class of membrane proteins that play important roles in nutrient scavenging, cell envelope assembly, protein folding, environmental signalling, host cell adhesion and virulence (Hutchings *et al.*, 2009; Paradis-Bleau *et al.*, 2010; Typas *et al.*, 2010). Precursor lipoproteins are translocated across the cytoplasmic membrane by either the Sec (general secretory) or Tat (twin arginine translocation) pathways (Thompson *et al.*, 2010) and then lipid modified by the enzyme lipoprotein diacylglycerol transferase (Lgt) on the sulphydryl group of a conserved cysteine. This cysteine is located in a lipobox motif, typically $L_{-3} - (A/S)_{-2} - (G/A)_{-1} - C_{+1}$, at the end of the signal peptide and becomes the N-terminal residue of the mature lipoprotein following cleavage of the signal peptide by lipoprotein signal peptidase (Lsp) (Hutchings *et al.*, 2009). In Gram-negative bacteria and Gram-positive mycobacteria a third enzyme, lipoprotein *N*-acyl transferase (Lnt), adds another fatty acid to the amino group of the conserved cysteine (Tschumi *et al.*, 2009; Vidal-Ingigliardi *et al.*, 2007). There is good evidence that *N*-acylation is involved in targeting lipoproteins to the outer membrane in *E. coli* and other Gram-negative bacteria (Robichon *et al.*, 2005). However, although mycobacteria have an outer membrane mycolic acid bilayer (Niederweis *et al.*, 2010; Sutcliffe, 2010; Zuber *et al.*, 2008) and are known to localise lipoproteins to the cell surface (Mawuenyega *et al.*, 2005; Niederweis *et al.*, 2010; Wu *et al.*, 1998), no mechanisms for the translocation of specific lipoproteins to the mycolate layer are known.

We recently characterised lipoprotein biogenesis in the model Gram-positive actinomycete bacterium *Streptomyces coelicolor* and found evidence that the pathway is essential in this organism. *S. coelicolor* is unusual in encoding two Lgt homologues and we were unable to construct a double *lgt* mutant. Furthermore, deletion of *lsp* seemingly resulted in the generation of secondary mutations that likely suppress the lethal phenotype of this mutation (Thompson *et al.*, 2010). To further investigate lipoprotein biogenesis in *Streptomyces* and in particular the role of Lgt and Lnt, we turned to the plant pathogenic species *Streptomyces scabies* which encodes single Lgt and Lsp enzymes. *S. scabies* 87-22 and other genome sequenced *Streptomyces* species also encode two Lnt homologues despite only containing a single lipid bilayer. *S. scabies* is a soil dwelling saprophytic bacterium with a complex developmental life cycle that has acquired a set of genes which enables it to colonise a variety of plants. Although best known as the causative agent of potato scab disease, *S. scabies* appears to be neither host, nor tissue, specific and will infect seedlings of monocotyledonous or dicotyledonous plants (Bignell *et al.*, 2010a). The main phenotype of an infected potato

plant is the presence of lesions or scabs on the surface of the potato, which can overlap to cover a significant proportion of its surface. These lesions are usually identified at the time of harvesting and they reduce the market value of the potato, making potato scab disease of great economic importance to farmers worldwide.

The role of lipoproteins in host cell attachment and virulence have been well studied in Gram-positive human and animal pathogens (Hutchings *et al.*, 2009; Sander *et al.*, 2004) but have only been studied from a bioinformatic perspective in plant pathogenic bacteria (Kovacs-Simon *et al.*, 2011; Sutcliffe & Hutchings, 2007). *Streptomyces* species export large numbers of lipoproteins via the Tat pathway (Shruthi *et al.*, 2010; Thompson *et al.*, 2010) and the Tat machinery is known to be essential for virulence in *S. scabies* (Joshi *et al.*, 2010). To understand the role of lipoproteins in the virulence of *S. scabies* and to further characterise lipoprotein biogenesis and the role of Lgt and Lnt in the genus *Streptomyces*, we identified putative Sec- and Tat-dependent lipoproteins in *S. scabies* and disrupted the lipoprotein biogenesis pathway by deleting the single *lgt* and *lsp* genes and the *lnt1* and *lnt2* genes individually and in combination. Here we report the effects of these mutations on growth and development, lipoprotein biogenesis and virulence in *S. scabies.* This in depth analysis of lipoprotein biogenesis is the first characterisation of Lgt and Lnt function in *Streptomyces* and, to our knowledge, the first study of lipoprotein biogenesis as a potential virulence determinant in a plant pathogen.

EXPERIMENTAL PROCEDURES

Bioinformatic identification of lipoproteins. All *S. scabies* proteins in StrepDB [http://strepdb.streptomyces.org.uk] were analysed for Sec- or Tat-dependent lipoprotein signal peptides as described previously (Thompson *et al.*, 2010). Functional predictions were based on the sequence annotation of these proteins, combined with BlastP analysis and conserved domain analyses.

Strains, plasmids, cosmids, primers and growth conditions. The strains, plasmids, cosmids and primers used in this study are listed in Table 1. *E. coli* was routinely grown in LB, or modified LB lacking NaCl to select for hygromycin resistance. *S. scabies* strains were grown on instant potato mash (IPM) agar (20g 'Smash' brand instant potato, 20g agar in 1 litre of tap water), Difco nutrient broth (DNB) or DNB agar (BD Diagnostics) and a 50:50 mix of

Tryptone Soya Broth (TSB; Oxoid) / Yeast extract malt extract (YEME) agar. Liquid cultures were grown in DNB or a 50:50 mix of TSB/YEME (Kieser, 2000).

Mutagenesis. Genes were replaced with apramycin resistance cassettes by PCR targeting the relevant cosmids from an ordered *S. scabies* cosmid library (Table 1) in *E. coli* using the Redirect method (Gust *et al.*, 2003). The primers used to construct and confirm the knockout cosmids and strains are listed in Table 1. All mutant and wild-type cosmids were checked by restriction digest and PCR with the relevant test primers before being introduced into *E. coli* strain ET12567/pUZ8002 and conjugated into *S. scabies* 87-22 as described previously (Joshi *et al.*, 2010). Mutant colonies were selected using the appropriate antibiotic and the mutation was checked by PCR using isolated genomic DNA from the mutant compared to the wild-type strain. Strains were complemented *in trans* in single copy by introduction of the genes and their native promoters on the integrative vector pSET152 in which the apramycin resistance cassette is replaced with a hygromycin resistance cassette. The gene encoding SCO3484-His was introduced on the same pSET152 vector (Table 1).

Microscopy. Brightfield images were acquired using a Zeiss M2 Bio Quad SV11 stereomicroscope. The samples were illuminated with a halogen lamp and reflected-light images captured with an AxioCam HRc CCD camera and AxioVision software (Carl Zeiss, Welwyn Garden City, UK). Preparation and imaging of samples using the scanning electron microscope was performed exactly as described (Thompson *et al.*, 2010). Images were saved as digital TIFF files and manipulated using Adobe Photoshop CS4.

Sample preparation for 2D gels. Cellophane discs were placed on YEME agar plates and inoculated with 10^5 spores or 10^6 spores of the Δ*lsp* and Δ*lgt* strains to give confluent growth after 48 hours. The mycelium was harvested and the lipoprotein fraction was extracted using Triton X-114 and prepared for 2D gel analysis as described previously. Extracellular protein samples were prepared as described previously (Thompson *et al.*, 2010).

2D PAGE. Analytical 2D page was carried out as described previously (Widdick *et al.*, 2006). For phenotypic analysis smaller 2D PAGE was used. Briefly, proteins were loaded onto 7cm pH4-7 IEF strips (GE Healthcare #17-6001-10) by rehydration at 20^{o}C for 12-14 hours then run using the following program: hold at 300V for 30 minutes, gradient to 1000V for 30 minutes, gradient to 5000V for 90 minutes then hold at 5000V for 5 hours. The 2nd dimension

Table 1. Strains, plasmids and primers

Strain		Reference
S. scabies		
87-22	Wild-type *S. scabies*	
BJT1040	*S. scabies Δlgt::apr*	This work
BJT1041	*S. scabies Δlgt::apr + lgt in trans*	This work
BJT1044	*S. scabies Δlsp::apr*	This work
BJT1045	*S. scabies Δlsp::apr + lsp in trans*	This work
BJT1047	*S. scabies Δlnt1::apr*	This work
BJT1048	*S. scabies Δlnt2::apr*	This work
BJT1049	*S. scabies Δlnt2::hyg*	This work
BJT1050	*S. scabies Δlnt1::apr Δlnt2::hyg*	This work
BW25113 (pIJ790)	BW25113 containing λ RED recombination	(Gust *et al.*, 2003)
ET12567 (pUZ8002)	*dam dcm* strain containing helper plasmid pUZ8002	(Gust *et al.*, 2003)
Plasmids		
pMS82	Integrative *Streptomyces* vector	(Gregory *et al.*, 2003)
pIJ10706	pSET152 with hygromycin resistance marker	(Kieser, 2000)
pBT110	pIJ10706 + full length *Streptomyces coelicolor lsp*	This work
pBT119	pMS82 + full length *Streptomyces coelicolor lgt1*	(Thompson *et al.*, 2010)
pBT120	pMS82 + truncated *Streptomyces coelicolor lgt1*	This work
pBT121	pMS82 + full length *Streptomyces coelicolor lgt2*	(Thompson *et al.*, 2010)
pBT123	pMS82 carrying the *S. scabies lsp* gene and its native promoter	This work
pBT124	pMS82 carrying the *S. scabies lgt* gene and its native promoter	This work
pAU3-45	pSET152 with thiostrepton resistance marker	(Bignell *et al.*, 2005; Thompson *et al.*, 2010)
pDWU59	pAU3-45 + SCO3484 6xHis	This work
Cosmids		
Scab 139	Cosmid with full length *S. scabies lgt*	
Scab 139 *lgt::apr*	Cosmid 139 with the *lgt* gene replaced with the *apr* cassette	
Scab 45	Cosmid + full length Scabies *lsp*	This work
Scab 45 *lsp::apr*	Cosmid 45 with the *lsp* gene replaced with the *apr* cassette	This work
Scab 351	Cosmid + full length Scabies *lnt1*	This work
Scab 351 *lnt1::apr*	Cosmid 351 with the *lnt1* gene replaced with the *apr* cassette	This work

Scab 2255	Cosmid + full length Scabies *lnt2*	This work
Scab 2255 *lnt2::hyg*	Cosmid 2255 with the *lnt2* gene replaced with the *hyg* cassette	This work
Scab 2255 *lnt1::apr*	See above	This work
Primers		
lgt KOfor	gtcggccgggcccgacagggtagcgtcgaccctgccatgattccggggatccg tcgacc	This work
lgt KOrev	tgggcaggcatcacatcgagcgaaccggctgtcgcgtcatgtaggctggagct gcttc	This work
lgt C-trunc rev	tcacggctccaccacggcctc	This work
lgt TESTfor	ggcggcagcggctcccgcaaggccg	This work
lgt TESTrev	tgggcaggcatcacatcgagcgaac	This work
Scab lgt comp For	acatctacgcggaccagtccatgac	This work
Scab lgt comp Rev	tgggcaggcatcacatcgagcgaac	This work
lnt1 KOfor	tgtcggcgccggggtgacgcgccgtacactccccgagtgattccggggatcc gtcgacc	This work
lnt1 KOrev	ccccgccgttccggcacgcgcacgcgaggcgtacggctatgtaggctggag ctgcttc	This work
lnt1 TESTfor	gatcagcgccgagctccccacggtg	This work
lnt1 TESTrev	ccccgccgttccggcacgcgcacgc	This work
lnt2 KOfor	actggagggacgcaccgggtggcagggaggcgctgtatgattccggggatc cgtcgacc	This work
lnt2 KOrev	gggcagggtgctcagcgcccgggacgccctggccactcatgtaggctggag ctgcttc	This work
lnt2 TESTfor	actggagggacgcaccgggt	This work
lnt2 TESTrev	gggcagggtgctcagcgccc	This work
P1	attccggggatccgtcgacc	This work
P2	tgtaggctggagctgcttc	This work
Scab Nec1 test for	atgagcgcgaacggaagcccc	This work
Scab Nec1 test rev	ctactttctcgttatccatat	This work
agarase test for	gtggtcaaccgacgtgatctc	This work
agarase test rev	ctacacggcctgatacgtcct	This work

was run on 1mm IPG well NuPAGE® 4-12% Bis-Tris ZOOM^{TM} Gels (Invitrogen #NP0330BOX) at 200V for 40 minutes using MES SDS running buffer (Invitrogen #NP0002). The gels were stained by using a colloidal Coomassie blue stain and then scanned using a GS800 Calibrated Densitometer (BioRad).

Purification of lipoproteins from membrane fractions. Solid cultures were grown on top of cellophane discs, harvested and then re-suspended in 20mM NaH_2PO_4 pH7.4 500mM NaCl (0.5ml/8.5cm plate), supplemented with lysozyme (30 mg/ml) and Complete^{TM} EDTA-free protease inhibitors (Roche). Mycelium was lysed by sonication on ice, until visual assessment indicated lysis was complete. Supernatants were separated into soluble and membrane fractions by ultracentrifugation (40,000 rpm, Beckman Ti70 rotor, 90 minutes at 4ºC). The supernatant at this stage was routinely retained as the cytoplasmic fraction. The pellet was re-suspended in 20mM NaH_2PO_4 pH7.4 500mM NaCl, 1% Triton-X-100, plus and Complete^{TM} EDTA-free protease inhibitors (Roche) (0.5ml/8.5cm plate) and incubated with gentle agitation at 4°C for 60 minutes. The mixture was then ultracentrifuged again, as previously. The supernatant was applied to a 1ml Hi-Trap Ni column equilibrated in 20mM NaH_2PO_4 pH7.4 500mM NaCl, 0.1% (v/v) Triton-X-100 using an ÄKTA FPLC system at 1ml/min. The column was washed in 20mM NaH_2PO_4 pH7.4, 500mM NaCl, 20mM Imidazole, 0.1% (v/v) Triton-X-100 and bound proteins were eluted using a gradient of 20-300mM Imidazole over 30ml.

Mass spectrometry. Purified lipoproteins were prepared using the method of Ujihara *et al.* (Ujihara *et al.*, 2008) and analysed as described previously (Tschumi *et al.*, 2009).

Potato necrosis assay. A potato tuber of the cultivar Maris Piper was washed and peeled. The tuber was surface sterilised by immersion in 2% bleach for 5 minutes, with stirring. The potato was transferred using sterile forceps into a beaker containing 500 ml of sterile water in the laminar flow hood. The potato was rinsed for 2 minutes, and transferred to a second beaker of sterile water where it was washed as before. After washing was completed, the potato was lain on some sterile paper towels and cores were taken using a sturdy potato peeler (sterilised with ethanol). These cores were sliced into discs 0.5 cm thick using a sterile scalpel. The slices were placed onto sterile filter paper, prewetted with 2 ml of sterile water, in a petri dish. Four potato slices were placed in each dish. Plates of the *S. scabies* strains were prepared on IPM agar. Spores were plated in order to obtain confluent lawns of bacteria. The plates were incubated for 5 nights, and agar plugs (1 cm^2) were cut using a sterile scalpel. A plug was placed on each potato piece, spore side down. Uninoculated IPM agar was used as a negative control. The plates were sealed and incubated at 30 °C in the dark for either 48 hours or 1 week. The agar plugs were then removed and the potato slices were viewed under a microscope to investigate signs of necrosis.

Radish seedling infection assay. Radish seeds of the cultivar Scarlet Globe were soaked in 70% ethanol for 10 minutes, the ethanol removed, then soaked in 13% bleach for a further 10 minutes. The bleach was removed, and the seeds were washed copiously in sterile H_2O. The seeds were placed onto sterile filter paper in a Petri dish, pre-wetted with 2 ml sterile H_2O. The seeds were incubated in the dark at room temperature for 24 hours. Cultures of various *S. scabies* strains were set up by growing 1 x 10^6 spores in 10 ml of 50% TSB/YEME media, with the relevant antibiotics at 30°C for 20 hours. After this time had elapsed the mycelium was collected by centrifugation and washed twice in TSB. After the final wash the mycelium was resuspended in 1 ml of TSB broth, with 500 µl being spread over half a square Petri dish containing plant MS medium (1% agar, added sucrose). The remaining 500 µl of mycelium were placed in a 12 well cell culture plate, into which seedlings at the same stage of germination were immersed prior to being placed in the Petri dish. These dishes were sealed and stored vertically in an incubation chamber at 21°C with a day length of 12 hours.

RESULTS

Bioinformatic identification and analysis of S. scabies *lipoproteins.*

Putative Sec- and Tat-dependent lipoproteins were identified by bioinformatic analysis of the *S. scabies* proteome as described previously (Thompson *et al.*, 2010) and are listed in Table S1. A total of 231 putative lipoproteins were identified, indicating that lipoproteins represent at least 2% of the *S. scabies* proteome. Approximately 10% of these putative lipoproteins are confidently predicted to be Tat substrates, indicating that ca. 18% of the total Tat substrates identified by Joshi *et al.* (Joshi *et al.*, 2010) are likely lipoproteins, including five Tat lipoproteins validated in the agarase secretion assay (Table S1). As in *S. coelicolor* and other actinomycetes, putative lipoprotein functions (Table 2) include many substrate binding proteins (SBP) of ABC transport systems (Table S2), enzymes, lipoproteins involved in cell envelope sensing, redox and cell wall homeostasis remodelling processes, and a significant proportion (35%) of hypothetical proteins of unknown function (Table S3). With regard to cell wall processess, seven putative lipoproteins belong to the YkuD family of L,D transpeptidases involved in peptidoglycan remodelling. As these lipoproteins are also well represented in *S. coelicolor* (Thompson *et al.*, 2010) it may be that these putative transpeptidases are significant in *Streptomyces* growth and development. Potential lipoprotein

Table 2. Lipoprotein functions in *Streptomyces scabies*.

Function	Numbers	%	Examples	References
Solute binding proteins	102	44%	SCAB81041 spermidine/putrescine SBP	(Joshi *et al.*, 2010)
Putative enzymes	21	9%	SCAB20691 polysaccharide deacetylase	
Redox processes	9	4%	CtaC, Cytochrome *c* oxidase subunit 2	(Thompson *et al.*, 2010)
Signal transduction (two component accessory lipoproteins)	5	2%	LpqB, CseA	(Hoskisson & Hutchings, 2006; Hutchings, 2007)
Cell envelope processes	13	6%	SCAB10101 putative transpeptidase	(Thompson *et al.*, 2010)
Function unknown	81	35%		

virulence factors include several putative degradative enzymes such as peptidases/proteases, four polysaccharide deacetylases, an α-*N*-acetylglucosaminidase and an α-L-arabinofuranosidase (Table S3). Also consistent with the acquisition of host derived nutrients are the identification of many SBP for sugars and amino acids/peptides (Table S2). As in *S. coelicolor* (Thompson *et al.*, 2010) it is notable that a significant proportion of the putative sugar-binding SBP (notably all those of the COG1653 family, n=36) are located in incomplete ABC operons that lack genes encoding the cognate energising ATPases. These ABC systems are presumptively energised by the single MsiK orthologue (SCAB_50161), which is orthologous to the multifunctional ATPase MsiK of *Streptomyces reticuli* (Schlosser *et al.*, 1997) and the promiscuous *Bacillus subtilis* MsmX ATPase (Ferreira & Sa-Nogueira, 2010). Notably, the pathogenicity island of *S. scabies* encodes three lipoproteins, SCAB77271, SCAB77361 and SCAB77471, all of which are putative COG1653 family SBP (Table S2) (Huguet-Tapia *et al.*, 2011). In addition, the putative lipoprotein SBP and confirmed Tat substrate SCAB81041 has been implicated in *S. scabies* virulence by a previous study which demonstrated that a *Δscab81041* mutant has a moderate virulence phenotype (Joshi *et al.*, 2010).

Disrupting lipoprotein biogenesis in S. scabies.

The lipoprotein biogenesis enzymes were identified by BLAST searching the *S. scabies* genome with the Lgt and Lsp sequences from *S. coelicolor* (Thompson *et al.*, 2010) and the Lnt sequences from *Escherichia coli*, *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* (Fig. S1). In contrast to *S. coelicolor* and *Streptomyces clavuligerus*, but in common with most other bacteria, *S. scabies* encodes a single Lgt enzyme (SCAB68531) along with a single Lsp enzyme (SCAB68121). The two putative *S. scabies* Lnt enzymes, which we have named Lnt1 (SCAB83111) and Lnt2 (SCAB76621), share a low amino acid sequence identity to each other (23%) and low levels of identity to *E. coli* Lnt (20% and 22%, respectively) and to *M. tuberculosis* Ppm1 (44% and 26%, respectively), which has been shown to have an N-terminal Lnt domain (Tschumi *et al.*, 2009). However, both Lnt1 and Lnt2 contain the three conserved residues (E267, K355, C387) that form the catalytic triad in *E. coli* Lnt and the mycobacterial Lnt enzymes (Fig. S1) (Buddelmeijer & Young, 2010; Vidal-Ingigliardi *et al.*, 2007). This led us to hypothesise that both Lnt homologues might catalyse the *N*-acylation of lipoproteins in *Streptomyces* because, to our knowledge, all genome-sequenced *Streptomyces* species encode both Lnt1 and Lnt2 homologues. An ordered cosmid library was used to replace each gene with an apramycin or hygromycin resistance cassette and these disrupted cosmids were conjugated into wild-type *S. scabies* strain 87-22. Single *lgt*, *lsp* and *lnt* mutants and a double *lnt* mutant were selected on the basis of apramycin and hygromycin resistance (see Experimental Procedures). Deleting *lsp* and *lgt* had a severe effect on the colony morphology of *S. scabies*, with both mutants exhibiting flatter, pock marked colonies with less obvious sporulation and the *lsp* null mutant in particular forming smaller colonies about half the size of the wild-type colonies. Deletion of the putative *lnt* genes individually or in combination had no visible effect on growth or development (Fig. 1A). The macroscopic phenotype of the *S. scabies lsp* mutant is less severe than that of an *S. coelicolor lsp* mutant but is consistent in that both form smaller, flatter colonies than their respective wild type strains. The *S. scabies lgt* and *lsp* mutants are also affected in development, they formed fewer coiled spore chains than the wild type and there is clearly less septation occurring in both the *lgt* and *lsp* mutants (Fig. 1B). These data are consistent with the phenotype of an *S. coelicolor lsp* mutant (Thompson *et al.*, 2010). These data suggests that disrupting lipoprotein biogenesis in *Streptomyces* at either the lipidation or signal cleavage stages has a serious effect on growth and development. The wild-type colony morphology was only partially restored in the *lgt* and *lsp* mutants by complementation *in*

trans with the *S. scabies lgt* or *lsp* genes (Fig. 1A) which suggests that secondary mutations may have arisen in the chromosomes of these strains. Previous results for *S. coelicolor* also suggested that disrupting *lsp* results in secondary mutations elsewhere in the chromosome (Thompson *et al.*, 2010). Heterologous expression of the *S. coelicolor lsp* gene also partially restored the wild-type colony morphology in the *S. scabies lsp* mutant as did both *S. coelicolor lgt1* and *lgt2* genes in the *S. scabies lgt* mutant. Significantly, this last result provides evidence that both *S. coelicolor lgt* genes encode functional Lgt homologues (Fig. 1A).

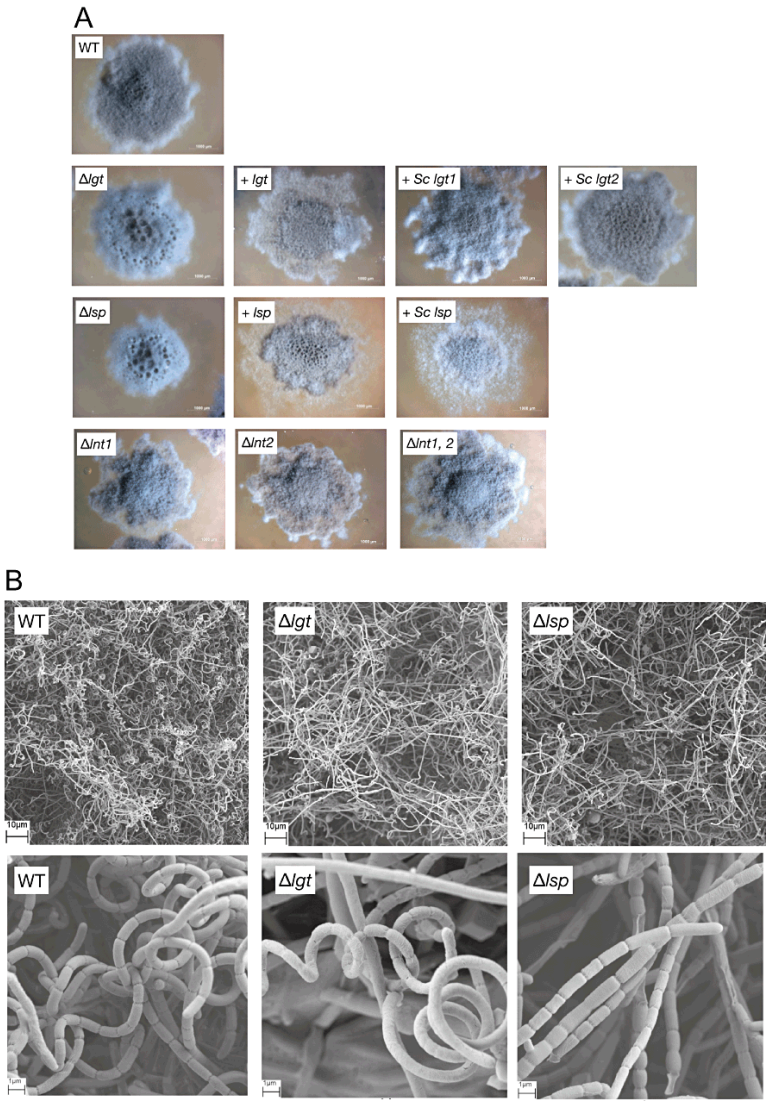


Fig. 1. A. Light microscope images (x 40 magnification) of single colonies of *S. scabies* wild-type 87-22, *Δlgt*, *Δlsp* and *Δlnt* strains as indicated, after 5 days growth on IPM agar. The *Δlgt* and *Δlsp* mutants form small and flatter colonies compared with the characteristic colonies of wild-type 87-22 that are raised in the centre. Complementation of the *Δlgt* and *Δlsp* mutants in single copy *in trans* only partially restores the wild-type phenotype. B. Scanning Electron Microscope images of single colonies of *S. scabies* wild-type 87-22, *Δlgt* and *Δlsp*. After five nights growth the mutant strains have fewer spore chains than the wild-type, suggesting a developmental delay.

Analysing the lipoproteome of S. scabies.

Our previous study in *S. coelicolor* analysed the effects of an *lsp* mutation on the lipoproteome of *Streptomyces* in some detail and demonstrated that lipoproteins were either lost from the membrane or retained with their signal peptides intact (Thompson *et al.*, 2010). In this work we focussed our efforts on identifying lipoproteins in the wild-type and *Δlgt* strains of *S. scabies*. Whole cell lysates of agar plate grown cultures were extracted with Triton X-114 and proteins in the lipophilic (detergent) fraction were precipitated with chloroform and methanol and resolved by two-dimensional (2D) gel electrophoresis, as described previously (Thompson *et al.*, 2010; Widdick *et al.*, 2006). MALDI TOF analysis of protein spots present in the wild-type samples revealed that 17 (~7.5%) of the putative lipoproteins identified by the bioinformatic analysis were present in the lipophilic fraction of the wild type strain, suggesting they are indeed lipoproteins (Table 3); Fig. 2A). No lipoproteins were detected in the lipophilic fraction of the *Δlgt* strain consistent with the idea that removing Lgt prevents lipidation of lipoproteins in *S. scabies* (Fig. 2B). Notably, the four proteins detected to be more abundant in this fraction are non-lipoproteins. A previous study on the role of Lgt in the Firmicute bacterium *Listeria* revealed that large numbers of lipoproteins are released into the growth medium in a *Δlgt* strain (Baumgärtner *et al.*, 2006). In *S. scabies* only 5 of the 17 lipoproteins identified in the wild-type lipophilic fraction were present in the extracellular fraction of the *Δlgt* strain along with another lipoprotein, SCAB19841, that was not detected in the wild-type lipophilic or cell wall wash fractions (Fig. S2, Tables 2 and S4). The remainder of the lipoproteins detected in the wild-type lipophilic fraction were not detected in either the lipophilic or cell wall wash fractions of the *lgt* strain, suggesting they are either not expressed (loss of SBPs would affect uptake of compounds and might affect regulation of cognate pathways) or that they are degraded by intramembrane or extracellular proteases. Loss of an unlipidated precursor of the CseA lipoprotein of *S. coelicolor* due to apparent proteolysis has been observed previously. The N-terminal peptides of the putative lipoproteins in the *lgt* cell wall wash fraction were not detected by MALDI TOF so conclusions cannot be drawn about the processing of lipoproteins shed into the extracellular environment in these strains.

For more rapid phenotypic comparison of the lipophilic fractions of the wild-type, *lgt*, *lsp* and *lnt* mutants and complemented mutant strains, samples were applied to 7 cm isoelectric focussing strips and separated in the second dimension on precast mini gels (see Experimental Procedures). These gels were run with biological duplicate samples and are for comparison

Table 3. Lipoproteins identified in wild-type *Streptomyces scabies* 87-22 and the isogenic *lgt* mutant by 2D PAGE.

Spot ID	Gene number	Function	Identified
1	SCAB05351	Putative enzyme (polyketide cyclase)	WT lipophilic fraction
2	SCAB08741	Putative SBP for peptides	WT lipophilic fraction
3	SCAB19051	Putative SBP for sugars	WT lipophilic fraction
4	SCAB21021, XylF	Putative SBP for pentose sugars	WT lipophilic fraction
5	SCAB24891	Putative SBP for amino acids	WT lipophilic fraction
6	SCAB27411	Putative SBP for oligopeptides	WT lipophilic fraction
7	SCAB29881	Putative SBP for amino acids	WT lipophilic fraction
8	SCAB29961	Putative SBP for sugars	WT lipophilic fraction
9	SCAB31531, BldKB	Putative SBP for peptides	WT lipophilic fraction
10	SCAB34981	Putative SBP for ribonucleosides	WT lipophilic fraction
11	SCAB49311, PstS	Putative SBP for phosphate.	WT lipophilic fraction
12	SCAB57751	Putative SBP for sugars	WT lipophilic fraction
13	SCAB68931	Putative SBP for branched chain amino acids	WT lipophilic fraction
14	SCAB73351	Putative SBP for dipeptides	WT lipophilic fraction
15	SCAB75441	Putative SBP for Fe^{3+}	WT lipophilic fraction
16	SCAB76561	Putative SBP	WT lipophilic fraction
17	SCAB85451	Putative SBP for Fe^{3+} hydroxamate	WT lipophilic fraction
18	SCAB19841	Putative SBP for nitrate-sulfonate-bicarbonate	*Δlgt* lipophilic fraction

only, no protein spots were analysed. In agreement with the 2D gel analysis reported above (Fig. 2) and previously (Thompson *et al.*, 2010) the protein patterns are markedly different between wild-type, *lgt* and *lsp* strains suggesting a loss or movement of lipoproteins from the lipophilic fraction in both the *lgt* and *lsp* mutants (representative gels are shown in Fig. 3). In contrast, the *lnt* single and double mutants were indistinguishable from the wild-type suggesting deletion of the *lnt* genes individually or in combination does not lead to the loss of lipoproteins from the lipophilic fraction (Fig. 3). These data are consistent with the observation that diacylated lipoproteins (generated by the action of Lgt and Lsp) are stably anchored to the membranes of monoderm bacteria. In the *lsp* mutant complemented with *S. coelicolor* or *S. scabies lsp* and in the *lgt* mutant complemented with either *S. scabies lgt* or with *S. coelicolor lgt1* or *lgt2* the pattern of protein spots in the lipophilic fraction was indistinguishable from the wild-type thus confirming that the loss or movement of protein spots in this fraction is due to loss of Lgt or Lsp activity (Figs. S3A and B). This also proves that both *S. coelicolor* Lgt1 (SCO2034) and Lgt2 (SCO7822) can lipidate lipoprotein

precursors and that *S. coelicolor* Lsp is functional in the heterologous host *S. scabies*. The Lgt1 and Lgt2 enzymes are 79% and 60% identical, respectively, to *S. scabies* Lgt while the *S. scabies* and *S. coelicolor* Lsp enzymes share 82% identity.

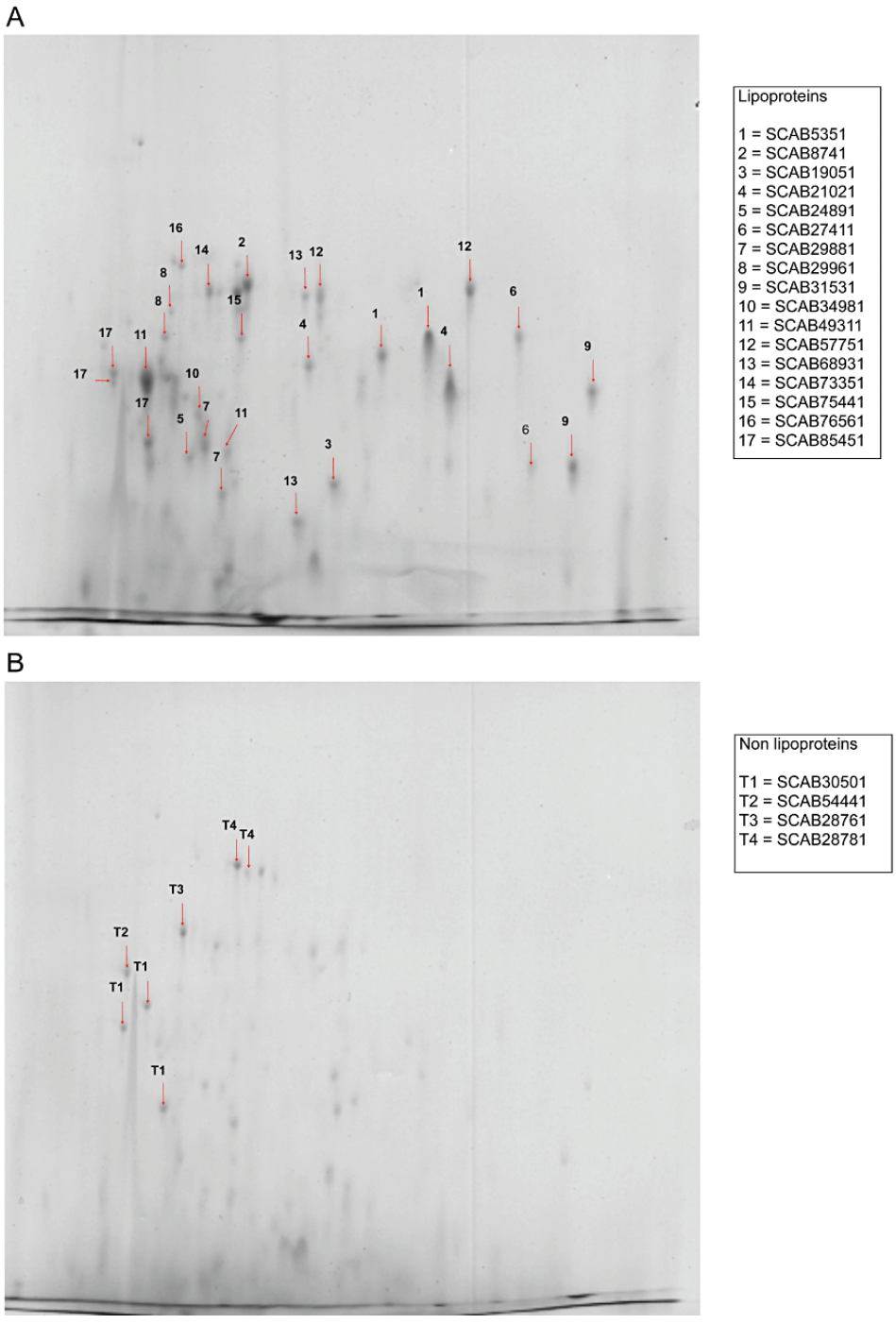


Fig. 2. Two dimensional polyacrylamide gels analysis of proteins extracted from the lipophilic fractions of the *S. scabies* wild-type (A) and *Δlgt* strains (B). Protein spots are highlighted with arrows and numbered with the corresponding proteins listed next to each gel. Functions of the putative lipoproteins are listed in Table 3. The four proteins identified as more abundant in the lipophilic fractions of the *Δlgt* strain are: the putative membrane protein SCAB30501 (T1); the putative enolase SCAB54441 (T2); the ATP synthase beta subunit SCAB28761 (T3); and the ATP synthase alpha subunit SCAB28781 (T4).

Lnt catalyses the N-*acylation of lipoproteins in* Streptomyces.

The results presented above suggest that neither Lnt homologue is required for correct lipoprotein localisation in the cytoplasmic membrane. However, this does not rule out a role for these putative Lnt enzymes in the *N*-acylation of lipoproteins in *Streptomyces* since diacylation by Lgt may be sufficient for membrane localisation. To determine whether Lnt1 or Lnt2 have lipoprotein *N*-acyl transferase activity we heterologously produced a His-tagged *S. coelicolor* lipoprotein, SCO3484, in the *S. scabies* wild-type and *lnt* mutant strains and used nickel affinity chromatography to purify SCO3484-His from membrane fractions isolated from each strain (see Experimental Procedures). The N-terminal modifications of SCO3484-His were examined using MALDI-TOF as described previously (Tschumi *et al.*, 2009). The predicted and experimentally determined m/z of the N-terminal peptides of SCO3484 with diacyl- or triacylation with C15, C16 and C17 fatty acids are shown in Table 3

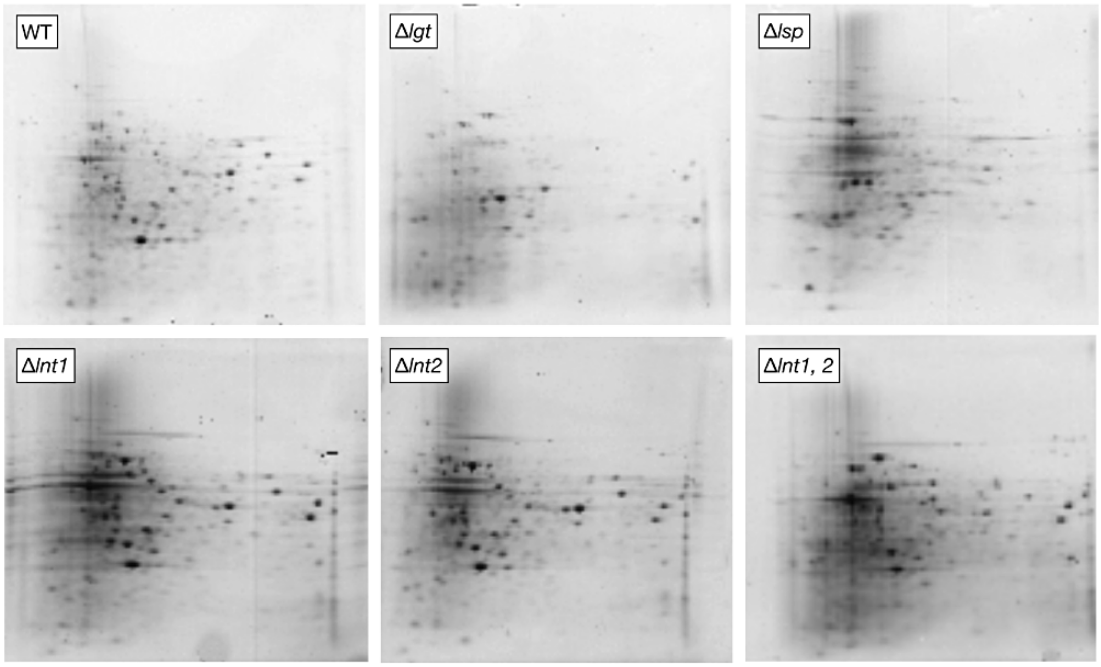


Fig. 3. Mini two-dimensional gel analysis of Triton X-114 extracted proteins from *S. scabies* wild-type 87-22, *Δlgt*, *Δlsp* and *Δlnt* strains after two days growth on YEME agar.

and indicate that the N-terminal peptide of SCO3484 is triacylated in wild-type *S. scabies* 87-22 (Table 3 and Fig. 4). The N-terminal fragment of SCO3484-His purified from the *lnt1* and *lnt1 lnt2* mutants is diacylated thus demonstrating that *lnt1* is required for *N*-acylation and that Lnt2 alone cannot compensate for the *lnt1* deletion in the *lnt1* single deletion. Conversely, the

SCO3484-His lipoprotein purified from the single *lnt2* mutant gives a mixture of diacylated and triacylated peptides (Table 3 and Fig. 4) indicating that although Lnt1 is sufficient for N-terminal acylation this step does not function as efficiently as in the wild type background. These data suggest either a subsidiary but inessential activity for Lnt2 or that Lnt2 does not

Table 3. Calculated m/z values and predicted chain lengths for diacylated and triacylated N-terminal tryptic peptides of the SCO3484-His lipoprotein purified from wild-type *S. scabies* 87-22, the isogenic *lgt* mutant complemented with *S. coelicolor lgt1* or *lgt2* and isogenic *lnt1*, *lnt2 and lnt1 lnt2* mutants as noted.

Calculated m/z	Predicted chain lengths	Observed m/z values of N-terminal Peptides					
		S. scabies wild-type	Δ*lgt* + *Sclgt1*	Δ*lgt* + *Sclgt2*	Δ*lnt1*	Δ*lnt2*	Δ*lnt1, 2*
2016.17	C15/C15/C15	2016.11	-	-	-	-	-
2030.19	C15/C15/C16	2030.14	2030.18	2030.18	-	2030.18	-
2044.17	C15/C15/C17	2044.15	2044.20	2044.20	-	2044.20	-
	C15/C16/C16						
2058.22	C15/C15/C18	2058.16	2058.22	2058.22	-	2058.22	-
	C15/C16/C17						
	C16/C16/C16						
2072.24	C15/C15/C19	2072.17	2072.22	-	-	-	-
	C15/C16/C18						
	C15/C17/C17						
	C16/C16/C17						
1791.17	C15/C15	-	-	-	-	-	1791.92
1805.99	C15/C16	-	-	-	1805.97	1805.97	1805.94
1820.01	C16/C16	-	-	-	1820.00	1820.02	1819.95
1834.02	C15/C18	-	-	-	1833.96	-	1833.98
	C16/C17						

function directly in lipoprotein *N*-acylation. In this respect it was previously observed that an *E. coli lnt* depletion strain could not be complemented with one of the *S. coelicolor lnt* homologues, *sco1336* (Vidal-Ingigliardi *et al.*, 2007). These data demonstrate that *S. scabies* Lnt1 is indeed a functional lipoprotein *N*-acyl transferase but also suggest that some *Streptomyces* Lnt homologues may not be functional in lipoprotein acylation. An alternative explanation may be a possible 'dominant negative' phenotype of *lnt1* deletion, that is, where deletion of *lnt1* affects Lnt2 function, although we cannot at present suggest a mechanism for this.

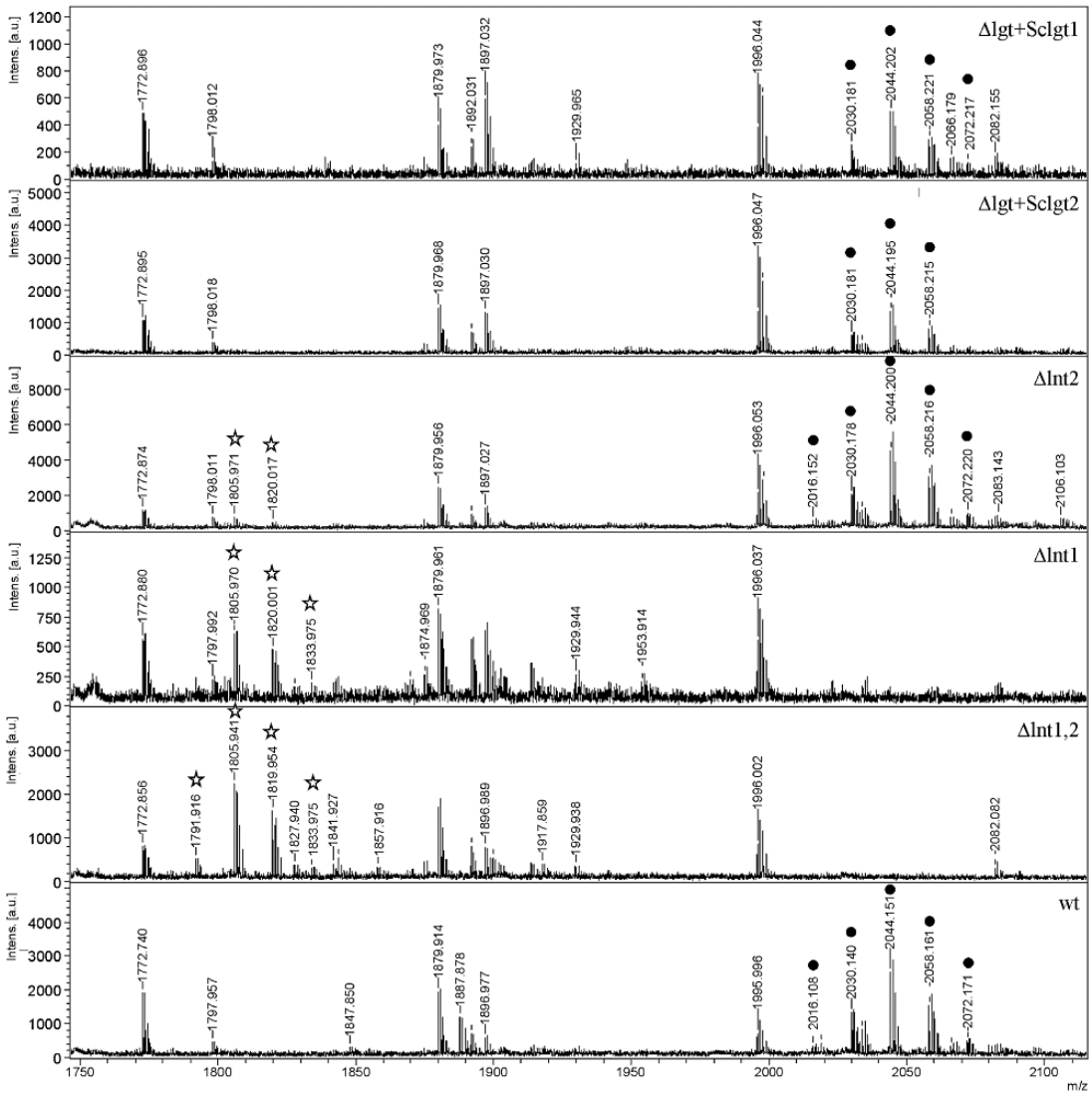


Fig 4. MALDI-TOF analysis with monoisotopic m/z values of SCO3484-His lipoprotein derived from the *S. scabies Δlgt* strain complemented with *S. coelicolor lgt1* or *lgt2* respectively, the *Δlnt1*, *Δlnt2* and *Δlnt1 lnt2* strains and the wild-type (wt) strain. Only m/z range from 1750-2100 is depicted. Filled circles: triacylated N-terminal peptides. Open stars: diacylated N-terminal peptides.

Lipoprotein biogenesis in Streptomyces *is ordered.*

The results presented above show that an *S. scabies lgt* mutant can be complemented by heterologously expressing either *lgt1* or *lgt2* from *S. coelicolor* and provides evidence that both enzymes encode functional Lgt enzymes (Figs 1A and 3). To analyse the N-terminal modifications catalysed by the two *S. coelicolor* Lgt enzymes, we purified SCO3484-His from the *S. scabies lgt* mutant complemented with either *S. coelicolor lgt1* or *lgt2*. In both

complemented strains we detected triacylated SCO3484-His suggesting that both enzymes diacylate this lipoprotein in an *S. scabies lgt* mutant and allow signal cleavage and *N*-acylation to occur (Table 3 and Fig. 4). We also purified SCO3484-His from the secreted fraction of the *S. scabies lgt* mutant and analysed the N-terminus by Edman sequencing.

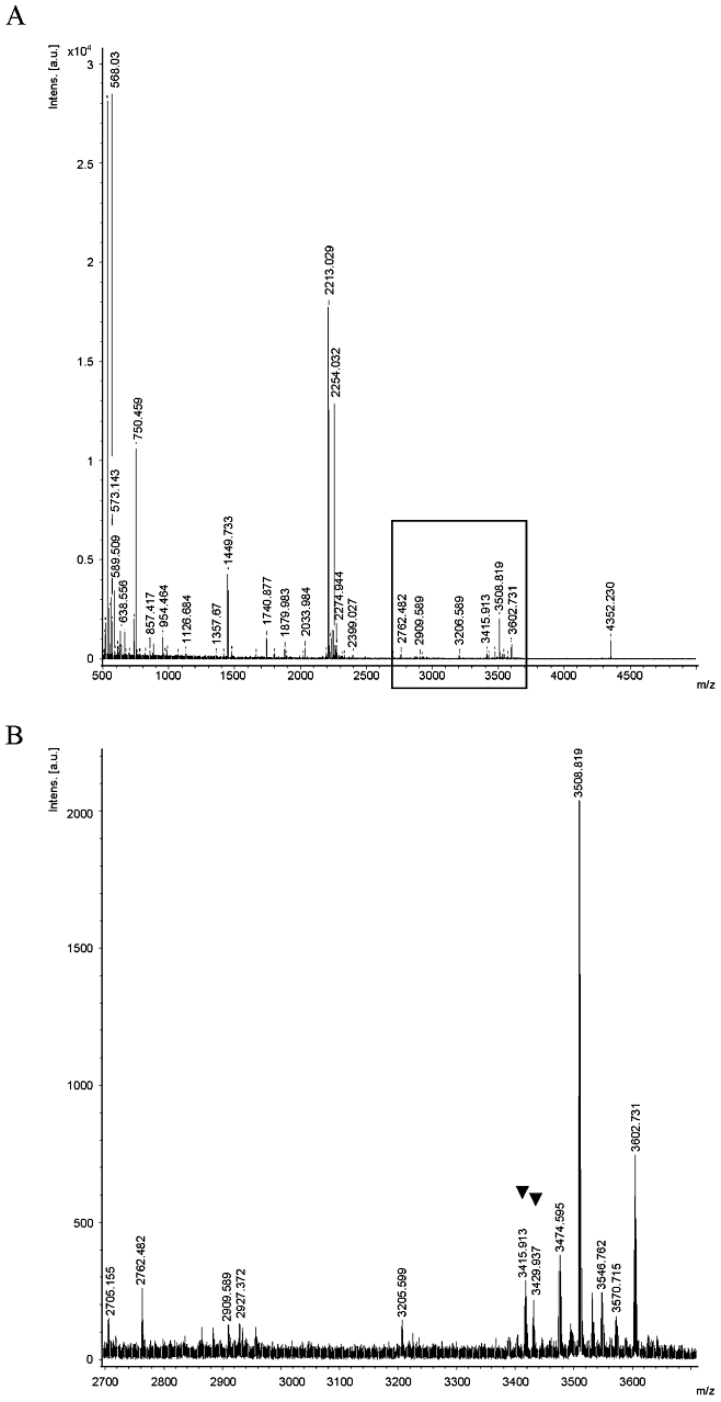


Fig 5. MALDI-TOF analysis with monoisotopic m/z values of SCO3484-His lipoprotein derived from the *S. scabies Δlsp* strain. Complete m/z range (A) and blown-up m/z range from 2700-3700 (B). Filled triangles show diacylated N-terminal peptides.

Although protein yield was low, sequencing of the N-terminus revealed a ragged N-terminus with the sequences starting with an alanine either from the -2 or -1 position, AAXSGGGNSS and AXSGGGNSSQ, respectively (data not shown), suggesting the signal sequence has most likely been cleaved by a Type I signal peptidase. This is consistent with the loss of Lgt function because Lsp rarely cleaves unlipidated substrates. Finally we analysed SCO3484-His purified from the membrane fraction of an *S. scabies lsp* mutant. In Gram-negative bacteria lipoprotein biogenesis occurs in strict order from Lgt to Lsp to Lnt, with Lsp only recognizing diacylated substrates. However, this is not true in all Gram-positive bacteria tested, since Lsp can cleave the signal peptide from unlipidated substrates in *Listeria* (Baumgärtner *et al.*, 2006). By definition, *N*-acylation can only occur after cleavage of the signal peptide because

Table 4. Calculated and predicted m/z values for the N-terminal tryptic peptide RGFLGGSLGVAGAVLLAACSGGGNSSQGSGGSK of the SCO3484-His lipoprotein purified from the isogenic lsp mutant

Calculated m/z	Predicted chain lengths	Observed m/z values
2879.44	-	n.f.
3415.93	C15/C16	3415.91
3429.95	C16/C16	3429.94

this event allows access to the amino terminus of the +1 cysteine. Analysis of the *S. scabies lsp* mutant identified an N-terminal fragment with the sequence RGFLGGSLGVAGAVLLAACSGGGNSSQGSGGSK whose m/z of 3414.88, 3415.91 and 3429.39 is consistent with the addition of C15/C16 or C16/C16 diacyl chains (Table 4 and Fig. 5). This fragment includes part of the uncleaved signal peptide of SCO3484 with a lipid group attached to the cysteine residue in the LAAC lipobox motif. These results show that lipidation occurs before signal peptide cleavage and that lipoprotein biogenesis in *Streptomyces* likely occurs in the order Lgt to Lsp to Lnt.

Disrupting lipoprotein biosynthesis has only a moderate effect on virulence.

To determine whether disrupting lipoprotein biogenesis has an effect on the virulence of *S. scabies*, the wild-type and mutant strains of *S. scabies* were first tested in a potato necrosis assay, whereby agar plugs containing each strain of *S. scabies* are placed face down onto

sterile potato slices, such that the bacteria are in direct contact with the potato, and an uninoculated agar plug is used as a negative control (Joshi *et al.*, 2007). This experiment was repeated at least ten times with each strain and taken together the data suggests that the *lgt*, *lsp* and *lnt* mutants cause the same levels of potato necrosis as the wild-type strain (representative images are shown in Fig. 6). To investigate the virulence of the wild-type and mutant strains on whole plants we used the previously established *S. scabies* radish seedling infection assay (Bignell *et al.*, 2010b). Sterilised seeds were germinated and then immersed in either uninoculated broth or broth inoculated with *S. coelicolor* (negative controls), *S. scabies* wild-type, *lgt*, *lsp* and *lnt* mutants (see Experimental Procedures). The seeds were plated onto plant MS medium with sucrose and grown for seven days. The negative control plant, overlaid with sterile growth medium, shows healthy growth, with a long primary root, and multiple secondary roots (Fig 7A). After 7 days the primary roots were typically between 11-12 cm in length. The leaves are healthy and new leaf growth can be seen at the apical meristem. The

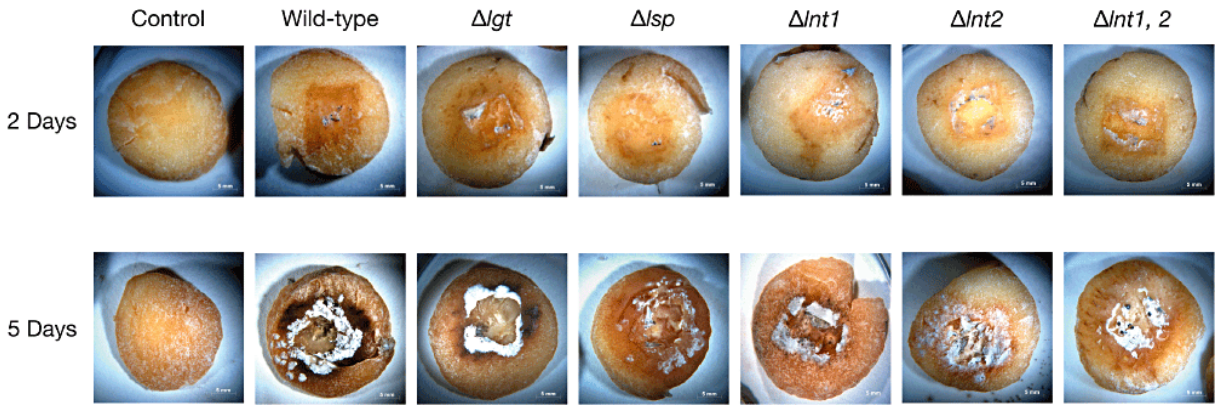


Fig 6. Sterile potato disks were exposed to uninoculated agar (control), or agar inoculated with wild-type *S. scabies* and the *S. scabies Δlgt*, *Δlsp*, *Δlnt1*, *Δlnt2* and *Δlnt1 lnt2* strains as shown and necrosis was examined after 2 days (top) and 5 days (bottom) growth. These are representative images from ten replicate experiments carried out for each strain.

stem of the plant was typically between 1.3 and 1.6 cm. This growth pattern is also seen in control plants treated with the non-pathogenic *S. coelicolor* (Fig. 7A). However, plants infected with wild-type *S. scabies* are stunted, with a primary root of between 4-6 cm and there is very little secondary root growth evident. The secondary roots that are present are much smaller than in the negative control (Fig. 7A). Very small leaves were also observed in the *S. scabies* wild-type infected plants, there was no evidence of further leaf growth and the stem was shorter than in the negative control plants, rarely exceeding 1.5 cm in length. This phenotype was relatively consistent for the wild-type and each of the *S. scabies* mutants tested

in ten separate experiments, although the main root was slightly longer in plants infected with the *lgt* or *lsp* mutants compared to wild-type and *lnt* mutants and contained more side branches, suggesting that disrupting lipoprotein biosynthesis might have a moderate effect on the virulence of *S. scabies* (Fig. 7A).

To further examine these effects the infected plant roots were examined using light microscopy (x 40 magnification). A complex network of root hairs can be seen in the plants

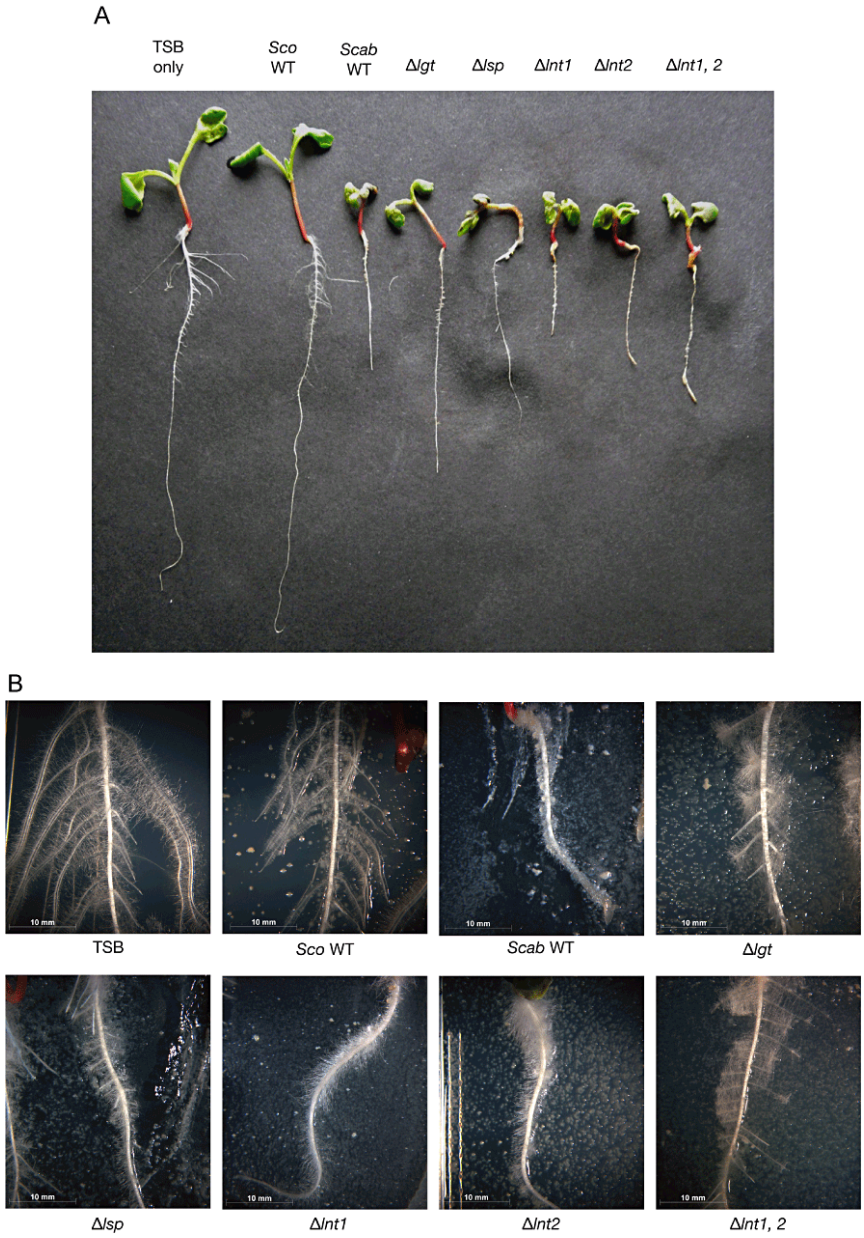


Fig 7. A. Radish seedlings were mock infected with growth medium (TSB) or wild-type *S. coelicolor* and with wild-type *S. scabies* and the isogenic *Δlgt*, *Δlsp, Δlnt1*, *Δlnt2* and *Δlnt1 lnt2* strains as shown, then grown for seven nights. All of the *S. scabies* infected plants were stunted in growth compared to the control plants although the *lgt* and *lsp* infected plants were consistently slightly larger than those infected with wild-type *S. scabies* suggesting a moderate effect on virulence. Assays were performed a minimum of ten times. B. Light microscope (x 40 magnification) examination of the roots of radish seedlings either mock infected with growth medium (TSB) or wild-type *S. coelicolor* and with wild-type *S. scabies* and the *Δlgt*, *Δlsp, Δlnt1*, *Δlnt2* and *Δlnt1 lnt2* strains as shown, after seven days growth.

treated with sterile growth medium or with growth medium inoculated with the non-pathogenic *S. coelicolor* (Fig. 7B). The entire root system grew into the agar making the plants difficult to remove without also removing a large amount of agar. In contrast, the root systems in plants infected with the *S. scabies* strains were much less complex. In the plants infected with wild-type *S. scabies* only the stunted primary root was present, with multiple brown nodules that represent failed secondary roots (Fig. 7B). These infected roots grow on the surface and cannot penetrate the growth medium. Only the fragile attachments of the root hairs allowed the plant to remain attached to the agar. This was also observed for the *S. scabies lgt*, *lsp* and *lnt* mutants. However, while infection with the *lgt*, *lsp* and *lnt* double mutants resulted in the stunted primary root seen in the wild-type infected plants, the overall phenotype is not as severe, and multiple secondary roots were consistently observed (Fig. 7B). Cumulatively, these data suggest a minor contribution of lipoproteins to the virulence of *S. scabies*, although since the *lgt* and *lsp* strains have growth and developmental phenotypes (Fig. 1) this reduced virulence may be indirect because these strains are less robust than the wild-type.

DISCUSSION

The results presented here have extended our previous work on *Streptomyces* lipoprotein biogenesis by investigating the function of *Streptomyces* Lgt and Lsp in the plant pathogen *S. scabies*. Deleting the *lsp* gene in *S. scabies* resulted in a very similar phenotype to the *S. coelicolor lsp* mutant, with a small colony phenotype and delayed sporulation that could not be fully complemented by reintroducing the *S. scabies lsp* gene. This supports the more in depth analysis of *lsp* in *S. coelicolor* which suggested that deletion of *lsp* results in secondary mutations that suppress the otherwise lethal effects of losing Lsp from the cell (Thompson *et al.*, 2010). Preliminary analysis of the genome sequences of the *S. coelicolor lsp* mutant and its parent strain suggest that a major transposition event has occurred in the *lsp* mutant (Widdick, Chandra and Hutchings, unpublished). *S. coelicolor* encodes two homologues of Lgt and the genes encoding these homologues could be deleted individually but not together in the same strain, which also suggested that lipoprotein biogenesis could be essential in this organism. In the current work the single *lgt* gene in *S. scabies* was successfully deleted and resulted in a growth and developmental phenotype similar to, but less severe than, that seen for the *S. coelicolor* and *S. scabies lsp* mutants. It is not clear why the *lsp* mutation has a more

severe effect than the *lgt* mutation in *S. scabies* but it is possibly due to the presence of lipoproteins with uncleaved signal sequences in the membrane which could accumulate and lead to membrane destabilisation and cell death. The *S. scabies lgt* mutation was only partially complemented by reintroduction of the *S. scabies lgt* gene, again supporting our suppressor mutation hypothesis. Two dimensional gel electrophoresis revealed a large scale loss of lipoproteins from the lipophilic fractions of the *lgt* mutant and a loss and major rearrangement of proteins in the *lsp* mutant suggesting the severe macro- and microscopic phenotypes observed in these strains result, at least partially, from loss of some or all of the lipoproteins from their membranes. Many of the lipoproteins missing from the *lgt* mutant membrane were detected in the extracellular fraction suggesting that, in the absence of lipidation, they are secreted out of the cell (Fig. 2 and Fig. S2). Lipoproteins were restored to the lipophilic fraction by *in trans* complementation with the corresponding genes from *S. coelicolor* or *S. scabies.* Crucially these results also confirm that *S. coelicolor* has two functional Lgt homologues, the first such demonstration for any bacterium.

Deletion of *lgt* or *lsp* had only a moderate effect on virulence in the radish seedling assay when compared with wild-type *S. scabies* and the uninfected controls (Fig. 7). The moderate effects on virulence may be due, at least in part, to the aberrant localisation or loss of the lipoprotein SCAB81041, removal of which has previously been demonstrated to moderately attenuate virulence of *S. scabies* (Joshi *et al.*, 2010). We also cannot discount the possibility that other lipoproteins may be involved in virulence. These could include the PAI lipoproteins SCAB77271, SCAB77361 and SCAB77471, although their functions and potential roles in the virulence of *S. scabies* remain to be investigated. None of these lipoproteins were detected in this work although it should be noted that the 2D gel analysis only identified ~10% of the predicted lipoproteome, consistent with similar experiments in *S. coelicolor* (Thompson *et al.*, 2010). However, it is also possible that the slower growing *lgt* and *lsp* mutant strains are less able to colonise the plant roots.

We have previously noted that actinomycete bacteria encode homologues of the Lnt enzymes which *N*-acylate lipoproteins in Gram-negative bacteria and direct them to the outer membrane via the Lol pathway (Hutchings *et al.*, 2009). It was subsequently shown that the *M. smegmatis* and *M. tuberculosis* Lnt homologues are true lipoprotein *N*-acyl transferase enzymes and are required for triacylation of lipoproteins in mycobacteria (Brülle *et al.*, 2010; Tschumi *et al.*, 2009). This finding is tantalising since mycobacteria have long been postulated to have an outer membrane made up of mycolic acids and this outer membrane bilayer hypothesis was recently proven experimentally (Hoffmann *et al.*, 2008; Zuber *et al.*,

2008). It is tempting therefore to hypothesise that lipoproteins may be transported from the cytoplasmic membrane to the outer membrane by an as yet undiscovered pathway. Such pathways must exist to transport proteins such as porins to the mycobacterial outer membrane and there is evidence that some lipoproteins in mycobacteria are surface exposed (Mawuenyega *et al.*, 2005; Niederweis *et al.*, 2010; Wu *et al.*, 1998). In this work we report that lipoproteins are triacylated in *Streptomyces* bacteria that apparently do not possess an outer membrane. This suggests that *N*-acylation of lipoproteins has an additional function in bacteria, or an alternative function in Gram-positive bacteria, and has not simply evolved to target lipoproteins to the Lol apparatus of Gram-negative bacteria. Intriguingly, *N*-acylation has also been reported in the low GC branch of Gram-positive bacteria Firmicutes despite a lack of Lnt homologues in the genomes of these bacteria (Kurokawa *et al.*, 2009; Navarre *et al.*, 1996). This again suggests there is a function for triacylation of lipoproteins in Gram-positive bacteria and in Gram-negative bacteria when they are retained in the cytoplasmic membrane by a sorting signal (Robichon *et al.*, 2005). No function for triacylation could be elucidated from this study because deletion of either *lnt1* or *lnt2* singly or in combination had no obvious effect on lipoprotein stability in the cytoplasmic membrane. An *lnt* double mutant was moderately affected in virulence, with infected roots looking similar to those infected by the *lgt* and *lsp* mutants (Fig. 7B). These data are difficult to interpret at present since an *lnt1* mutation has the same effect as an *lnt* double mutation in terms of lipidation of SCO3484 but the *lnt1* mutant was not affected in virulence (Fig. 7B).

Importantly this is the first study of a bacterium with more than one Lnt enzyme and our results indicate that Lnt1 is indeed a lipoprotein *N*-acyl transferase. Although Lnt2 was not demonstrated to compensate for loss of Lnt1, the data suggest that both Lnt enzymes are required for efficient *N*-acylation of the heterologously expressed *S. coelicolor* lipoprotein SCO3484 as removal of Lnt2 resulted in a mixture of diacylated and triacylated SCO3484 whereas only triacylated SCO3484 was detected in the wild type background. This implies Lnt2 does play a role in *N*-acylation of lipoproteins, perhaps through the formation of heterodimers with Lnt1. Despite sharing only 22% amino acid sequence identity both Lnt1 and Lnt2 contain the conserved catalytic residues identified in *E. coli* Lnt and so we do not exclude the possibility that Lnt2 alone can catalyse *N*-acylation of other lipoproteins in *S. scabies*. It will be necessary to analyse the N-terminal modifications of multiple native *S. scabies* lipoproteins in the wild-type and *lnt* strains to fully elucidate the functions of these enzymes but given the technical challenge of purifying lipoproteins expressed at native levels this will not be trivial. We conclude that while *Streptomyces* are the first monoderm bacteria

to contain experimentally verified Lnt enzymes, alternative enzymes unrelated to Lnt at the sequence level likely exist to catalyse *N*-acylation in the low GC Firmicute bacteria. The function of *N*-acylation in monoderm bacteria remains to be discovered.

In summary, this work has shown that lipoprotein biogenesis is likely a highly significant pathway in *S. scabies* as it is in *S. coelicolor*: although we were able to generate both *lgt* and *lsp* mutants these exhibited severe defects in growth and development, consistent with the bioinformatic analysis that lipoproteins are an abundant and functionally diverse category of protein in *S. scabies*. In addition to, or perhaps because of, their growth phenotypes, the *lgt* or *lsp* strains were also less virulent than wild-type *S. scabies.* This is also consistent with the identification of several candidate lipoproteins that could be involved in virulence and individual knockout mutations in these candidate lipoprotein genes will be necessary to elucidate their roles in plant infection. Crucially, this work has also extended the previous study by discovering a third *N*-acylation step in the *Streptomyces* lipoprotein biogenesis pathway. We have confirmed that *Streptomyces* encode functional Lnt enzymes that triacylate their lipoproteins. Future work will be aimed at elucidating the functions of Lnt1 and Lnt2 and mapping the secondary mutations that arise when lipoprotein biogenesis is disrupted in *Streptomyces* species.

ACKNOWLEDGEMENTS

We thank Rose Loria for providing *S. scabies* cosmids and Ramon Santa-Maria and Dawn Bignell for useful discussions. We also thank Paul Thomas for assistance with light microscopy in the Henry Wellcome Laboratory for Cell Imaging at UEA, Kim Findlay for the scanning electron microscopy, Fran Mullholland for technical advice and assistance with 2D gels, Mike Naldrett and Gerhardt Saalbach for MALDI-TOF analysis of 2D gel spots, Ryan Seipke for advice on plant infection assays and Yolanda Joho-Auchli from the functional Genomics Center Zurich for MALDI-TOF/TOF analysis of SCO3484. This work was funded by the Biotechnology and Biological Sciences Research Council (grant numbers F009224/1 and F009429/1 and a PhD studentship to BJT). M.I.H. also acknowledges support from a Research Councils UK Academic Fellowship and from UEA and the Royal Society for this work. Work in the laboratory of Peter Sander is supported by the Swiss National Science Foundation.

ADDENDUM

Personal contribution to chapter 4

My contribution as a co-author to this manuscript was as follows:

- MALDI-TOF data analysis
- MALDI-TOF/TOF data analysis
- Proof-reading of the manuscript

OUTLOOK

Lipidation of proteins is required for their anchoring into membranes and their sorting to the cell surface in all bacteria. Recent studies aimed at elucidating the role of lipoproteins and their synthesizing enzymes in the biology of mycobacteria. Lipoproteins are crucial for the synthesis of the mycobacterial cell wall, uptake of nutrients, signalling, sensing of environmental stress or host-pathogen interactions.

In the present study the membrane anchor of lipoproteins from *M. smegmatis*, *M. bovis* and *S. scabies* were characterized at the molecular level. Lipoproteins of these species are triacylated as they carry a diacylglyceryl residue and an additional *N*-acyl. Triacylation of lipoproteins therefore seems to be a conserved mechanism in both Gram-negative and G+C-rich Gram-positive bacteria.

In Gram-negative bacteria, *N*-acylation by Lnt is a prerequisite for the transport of lipoproteins to the outer membrane by the Lol system. Likewise, *N*-acylation of mycobacterial lipoproteins may be required for the transport to the mycobacterial outer membrane, which resembles the outer membrane of Gram-negative bacteria. But, there are no Lol homologues present in the genomes of mycobacteria and their transport system still remains under investigation. Possible approaches to identify the transport system could be the expression of a lipoprotein construct with a reporter suitable for screening transposon libraries of mycobacteria, or *in vivo* photocrosslinking, by which a putative transporter is bound to a reporter lipoprotein and afterwards purified and analysed by fingerprint.

By MALDI-TOF/TOF analyses, palmitic, oleic and tuberculostearic acid were identified as ester-linked fatty acids of the diacylglyceryl residue, while palmitic and tuberculostearic acid were found as *N*-acyl in mycobacteria. In *S. scabies*, C15 to C18 fatty acids were found to be ester-linked fatty acids of the diacylglyceryl residue, while C15 up to C19 fatty acids were found as *N*-acyl. Thus, the mycobacterial lipoprotein-synthezising enzymes differ in substrate specificity from that in *E. coli*. The recognition of lipoproteins by the innate host-immune system is mediated through the perception of the di- or triacylation by TLR-2 and its co-receptors TLR-1 and TLR-6, respectively. The effect of the acyl-chain length and the composition of the fatty acids on TLR signalling have not been investigated in detail. If the selection and arrangement of fatty acids of the membrane anchor is crucial for the function of the lipoprotein, its recognition by transporters or for its correct localization in bacteria is another interesting point for further investigation.

Beside the modifications with fatty acids, lipoproteins LprF and LppX in this study were also found with hexose glycosylations. *O*-glycosylations are known to occur at Ser and Thr residues, but the exact glycosylation site within the peptide is still unknown. Although an

influence of glycosylation on the correct subcellular lipoprotein localization and its protection from proteolytic degradation is proposed, the information about its function isscarce. Generation of deletion mutants lacking the polyprenol-monophosphomannose synthase (Ppm) gene in *M. smegmatis* would be one option to investigate the effects of glycosylation on lipoprotein function, localization or degradation. Of note, Ppm is essential in *M. tuberculosis*. A tool for the identification of glycosylation sites could be lipoprotein analyses with mass spectrometry in combination with site directed mutagenesis.

Within this study, BCG_2070c was confirmed as functional Lnt in *M. bovis* BCG. But, a second open reading frame, BCG_2279c homologous to *E. coli* Lnt was found. *Streptomyces ssp.* also encode two Lnt homologues, which both seem to be required for efficient lipoprotein *N*-acylation. Although BCG_2279c was not able to compensate Lnt function in the BCG_2070c deletion mutant in this study, a role of BCG_2279c as Lnt under special conditions could be supposed. Therefore, analyses of lipoproteins from BCG_2070c deletion mutants grown under specific growth conditions like stress or nutrient starvation would be an approach to investigate the role of the second Lnt homologue.

The antituberculosis drug pipeline is still not sufficiently filled and the vaccines used at present do not provide effective protection against tuberculosis in adults. Due to their contribution to virulence, lipoproteins and their synthesizing enzymes are promising candidates as drug targets. Moreover, some lipoproteins confer a protective immune response and may therefore be used as subunit vaccines. But, despite progress in lipoprotein research much remains to be learned with respect to the synthesis, function and localization of mycobacterial lipoproteins and their role in host-pathogen interactions. The results of this study may contribute to provide the knowledge which is required on the way to combat tuberculosis.

REFERENCES

Abubakar, I., Zignol, M., Falzon, D. & other authors (2013). Drug-resistant tuberculosis: time for visionary political leadership. *The Lancet infectious diseases* **13**, 529-539.

Almeida Da Silva, P. E. & Palomino, J. C. (2011). Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. *The Journal of antimicrobial chemotherapy* **66**, 1417-1430.

Anwari, K., Poggio, S., Perry, A. & other authors (2010). A modular BAM complex in the outer membrane of the alpha-proteobacterium Caulobacter crescentus. *PLoS One* **5**, e8619.

Asanuma, M., Kurokawa, K., Ichikawa, R., Ryu, K. H., Chae, J. H., Dohmae, N., Lee, B. L. & Nakayama, H. (2011). Structural evidence of alpha-aminoacylated lipoproteins of Staphylococcus aureus. *FEBS J* **278**, 716-728.

Azad, A. K., Sadee, W. & Schlesinger, L. S. (2012). Innate immune gene polymorphisms in tuberculosis. *Infect Immun* **80**, 3343-3359.

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L. & Sankaran, K. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* **188**, 2761-2773.

Baena, A. & Porcelli, S. A. (2009). Evasion and subversion of antigen presentation by Mycobacterium tuberculosis. *Tissue Antigens* **74**, 189-204.

Banaiee, N., Kincaid, E. Z., Buchwald, U., Jacobs, W. R., Jr. & Ernst, J. D. (2006). Potent inhibition of macrophage responses to IFN-gamma by live virulent Mycobacterium tuberculosis is independent of mature mycobacterial lipoproteins but dependent on TLR2. *J Immunol* **176**, 3019-3027.

Banaiee, N., Jacobs, W. R. & Ernst, J. D. (2007). LspA-independent action of globomycin on Mycobacterium tuberculosis. *The Journal of antimicrobial chemotherapy* **60**, 414-416.

Baumgärtner, M., Kärst, U., Gerstel, B., Loessner, M., Wehland, J. & Jänsch, L. (2006). Inactivation of Lgt allows systematic characterization of lipoproteins from Listeria monocytogenes. *J Bacteriol* **189**, 313-324.

Bernadac, A., Gavioli, M., Lazzaroni, J. C., Raina, S. & Lloubes, R. (1998). Escherichia coli tol-pal mutants form outer membrane vesicles. *J Bacteriol* **180**, 4872-4878.

Bigi, F., Gioffre, A., Klepp, L. & other authors (2004). The knockout of the lprG-Rv1410 operon produces strong attenuation of Mycobacterium tuberculosis. *Microbes Infect* **6**, 182-187.

Bignell, D. R., Tahlan, K., Colvin, K. R., Jensen, S. E. & Leskiw, B. K. (2005). Expression of ccaR, encoding the positive activator of cephamycin C and clavulanic acid production in Streptomyces clavuligerus, is dependent on bldG. *Antimicrob Agents Chemother* **49**, 1529-1541.

Bignell, D. R., Huguet-Tapia, J. C., Joshi, M. V., Pettis, G. S. & Loria, R. (2010a). What does it take to be a plant pathogen: genomic insights from Streptomyces species. *Antonie Van Leeuwenhoek* **98**, 179-194.

Bignell, D. R., Seipke, R. F., Huguet-Tapia, J. C., Chambers, A. H., Parry, R. J. & Loria, R. (2010b). Streptomyces scabies 87-22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant-microbe interactions. *Mol Plant Microbe Interact* **23**, 161-175.

Bloom, B. R. & Murray, C. J. (1992). Tuberculosis: commentary on a reemergent killer. *Science* **257**, 1055-1064.

Bos, M. P., Tefsen, B., Geurtsen, J. & Tommassen, J. (2004). Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proc Natl Acad Sci U S A* **101**, 9417-9422.

Bos, M. P. & Tommassen, J. (2011). The LptD chaperone LptE is not directly involved in lipopolysaccharide transport in Neisseria meningitidis. *J Biol Chem* **286**, 28688-28696.

Böttger, E. C. & Springer, B. (2008). Tuberculosis: drug resistance, fitness, and strategies for global control. *European journal of pediatrics* **167**, 141-148.

Brennan, P. J. & Nikaido, H. (1995). The envelope of mycobacteria. *Annual review of biochemistry* **64**, 29-63.

Brennan, P. J. & Crick, D. C. (2007). The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery. *Current topics in medicinal chemistry* **7**, 475-488.

Brosch, R., Gordon, S. V., Garnier, T. & other authors (2007). Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* **104**, 5596-5601.

Brülle, J. K., Grau, T., Tschumi, A., Auchli, Y., Burri, R., Polsfuss, S., Keller, P. M., Hunziker, P. & Sander, P. (2010). Cloning, expression and characterization of Mycobacterium tuberculosis lipoprotein LprF. *Biochem Biophys Res Commun* **391**, 679-684.

Bubeck Wardenburg, J., Williams, W. A. & Missiakas, D. (2006). Host defenses against Staphylococcus aureus infection require recognition of bacterial lipoproteins. *Proc Natl Acad Sci U S A* **103**, 13831-13836.

Buddelmeijer, N. & Young, R. (2009). The essential Escherichia coli apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry* **49**, 341-346.

Buddelmeijer, N. & Young, R. (2010). The essential Escherichia coli apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry* **49**, 341-346.

Buwitt-Beckmann, U., Heine, H., Wiesmuller, K. H., Jung, G., Brock, R., Akira, S. & Ulmer, A. J. (2005). Toll-like receptor 6-independent signaling by diacylated lipopeptides. *European journal of immunology* **35**, 282-289.

Camacho, L. R., Ensergueix, D., Perez, E., Gicquel, B. & Guilhot, C. (1999). Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. *Mol Microbiol* **34**, 257-267.

Chao, M. C. & Rubin, E. J. (2010). Letting sleeping dos lie: does dormancy play a role in tuberculosis? *Annu Rev Microbiol* **64**, 293-311.

Chen, M. H., Takeda, S., Yamada, H., Ishii, Y., Yamashino, T. & Mizuno, T. (2001). Characterization of the RcsC-->YojN-->RcsB phosphorelay signaling pathway involved in capsular synthesis in Escherichia coli. *Bioscience, biotechnology, and biochemistry* **65**, 2364-2367.

Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. (2002). Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **99**, 11020-11024.

Chng, S. S., Ruiz, N., Chimalakonda, G., Silhavy, T. J. & Kahne, D. (2010). Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane. *Proc Natl Acad Sci U S A* **107**, 5363-5368.

Clavel, T., Germon, P., Vianney, A., Portalier, R. & Lazzaroni, J. C. (1998). TolB protein of Escherichia coli K-12 interacts with the outer membrane peptidoglycan-associated proteins Pal, Lpp and OmpA. *Mol Microbiol* **29**, 359-367.

Cole, S. T., Brosch, R., Parkhill, J. & other authors (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537-544.

Cole, S. T. & Riccardi, G. (2011). New tuberculosis drugs on the horizon. *Current opinion in microbiology* **14**, 570-576.

Cole, S. T., Eisenach, K.D., McMurray, D.N., and Jacobs, W.R. (2005). Tuberculosis and the Tubercle Bacillus. Washington, D.C.: ASM Press.

Cosma, C. L., Sherman, D. R. & Ramakrishnan, L. (2003). The secret lives of the pathogenic mycobacteria. *Annu Rev Microbiol* **57**, 641-676.

Daffe, M. & Draper, P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39**, 131-203.

Daffe, M. (2008). The global architecture of the mycobacterial cell envelope. In The Mycobacterial Cell Envelope, pp. 3-11. Edited by M. Daffe, and Reyrat, J.-M. Washington, D.C.: ASM Press.

Daniel, T. M. (2006). The history of tuberculosis. *Respiratory medicine* **100**, 1862-1870.

Davidson, A. L. & Chen, J. (2004). ATP-binding cassette transporters in bacteria. *Annual review of biochemistry* **73**, 241-268.

Dawson, R. J. & Locher, K. P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature* **443**, 180-185.

Dawson, R. J., Hollenstein, K. & Locher, K. P. (2007). Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Mol Microbiol* **65**, 250-257.

Dawson, R. J. & Locher, K. P. (2007). Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. *FEBS Lett* **581**, 935-938.

DeChavigny, A., Heacock, P. N. & Dowhan, W. (1991). Sequence and inactivation of the pss gene of Escherichia coli. Phosphatidylethanolamine may not be essential for cell viability. *J Biol Chem* **266**, 10710.

Desvaux, M., Hebraud, M., Talon, R. & Henderson, I. R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol* **17**, 139-145.

Dhiman, R. K., Dinadayala, P., Ryan, G. J., Lenaerts, A. J., Schenkel, A. R. & Crick, D. C. (2011). Lipoarabinomannan localization and abundance during growth of Mycobacterium smegmatis. *J Bacteriol* **193**, 5802-5809.

Diaz-Silvestre, H., Espinosa-Cueto, P., Sanchez-Gonzalez, A., Esparza-Ceron, M. A., Pereira-Suarez, A. L., Bernal-Fernandez, G., Espitia, C. & Mancilla, R. (2005). The 19-kDa antigen of Mycobacterium tuberculosis is a major adhesin that binds the mannose receptor of THP-1 monocytic cells and promotes phagocytosis of mycobacteria. *Microb Pathog* **39**, 97-107.

Dietrich, J. & Doherty, T. M. (2009). Interaction of Mycobacterium tuberculosis with the host: consequences for vaccine development. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **117**, 440-457.

Drage, M. G., Pecora, N. D., Hise, A. G., Febbraio, M., Silverstein, R. L., Golenbock, D. T., Boom, W. H. & Harding, C. V. (2009). TLR2 and its co-receptors determine responses of macrophages and dendritic cells to lipoproteins of Mycobacterium tuberculosis. *Cell Immunol* **258**, 29-37.

Drage, M. G., Tsai, H. C., Pecora, N. D. & other authors (2010). Mycobacterium tuberculosis lipoprotein LprG (Rv1411c) binds triacylated glycolipid agonists of Toll-like receptor 2. *Nature structural & molecular biology* **17**, 1088-1095.

Ducati, R. G., Ruffino-Netto, A., Basso, L. A. & Santos, D. S. (2006). The resumption of consumption -- a review on tuberculosis. *Memorias do Instituto Oswaldo Cruz* **101**, 697-714.

Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W. & Schlosser, A. (1998). The ABC maltose transporter. *Mol Microbiol* **29**, 685-694.

Ewald, S. E. & Barton, G. M. (2011). Nucleic acid sensing Toll-like receptors in autoimmunity. *Current opinion in immunology* **23**, 3-9.

Farrow, M. F. & Rubin, E. J. (2008). Function of a mycobacterial major facilitator superfamily pump requires a membrane-associated lipoprotein. *J Bacteriol* **190**, 1783-1791.

Ferreira, M. J. & Sa-Nogueira, I. (2010). A multitask ATPase serving different ABC-type sugar importers in Bacillus subtilis. *J Bacteriol* **192**, 5312-5318.

Foster, T. J. (2005). Immune evasion by staphylococci. *Nat Rev Microbiol* **3**, 948-958.

Fraser, C. M., Casjens, S., Huang, W. M. & other authors (1997). Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. *Nature* **390**, 580-586.

Fukuda, A., Matsuyama, S., Hara, T., Nakayama, J., Nagasawa, H. & Tokuda, H. (2002). Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. *J Biol Chem* **277**, 43512-43518.

Gao, B., Paramanathan, R. & Gupta, R. S. (2006). Signature proteins that are distinctive characteristics of Actinobacteria and their subgroups. *Antonie Van Leeuwenhoek* **90**, 69-91.

Garbe, T., Harris, D., Vordermeier, M., Lathigra, R., Ivanyi, J. & Young, D. (1993). Expression of the Mycobacterium tuberculosis 19-kilodalton antigen in Mycobacterium smegmatis: immunological analysis and evidence of glycosylation. *Infect Immun* **61**, 260-267.

Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolynec, P. P., Likic, V. A., Purcell, A. W., Buchanan, S. K. & Lithgow, T. (2008). Protein secretion and outer membrane assembly in Alphaproteobacteria. *FEMS Microbiol Rev* **32**, 995-1009.

Gengenbacher, M. & Kaufmann, S. H. (2012). Mycobacterium tuberculosis: success through dormancy. *FEMS Microbiol Rev* **36**, 514-532.

Gennity, J. M. & Inouye, M. (1991). The protein sequence responsible for lipoprotein membrane localization in Escherichia coli exhibits remarkable specificity. *J Biol Chem* **266**, 16458-16464.

Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *The Journal of cell biology* **164**, 19-24.

George, K. M., Chatterjee, D., Gunawardana, G., Welty, D., Hayman, J., Lee, R. & Small, P. L. (1999). Mycolactone: a polyketide toxin from Mycobacterium ulcerans required for virulence. *Science* **283**, 854-857.

Gerken, H., Leiser, O. P., Bennion, D. & Misra, R. (2010). Involvement and necessity of the Cpx regulon in the event of aberrant beta-barrel outer membrane protein assembly. *Mol Microbiol* **75**, 1033-1046.

Gibbons, H. S., Wolschendorf, F., Abshire, M., Niederweis, M. & Braunstein, M. (2007). Identification of two Mycobacterium smegmatis lipoproteins exported by a SecA2-dependent pathway. *J Bacteriol* **189**, 5090-5100.

Gilleron, M., Jackson, M., Nigou, J., and Puzo, G. (2008). Structure, biosynthesis and activities of the phosphatidyl-myo-inositol-based lipoglycans. In The Mycobacterial Cell Envelope, pp. 75-105. Edited by M. Daffe, and Reyrat, J.- M. Washington D.C.: ASM press.

Goren, M. B., and Brennan, P.J. (1979). Mycobacterial lipids: chemistry and biological activities. In Tuberculosis, pp. 63-193. Edited by G. P. Youmans: The W. B. Saunders Co., Philadelphia, PA.

Gowthaman, U., Rai, P. K., Khan, N., Jackson, D. C. & Agrewala, J. N. (2012). Lipidated promiscuous peptides vaccine for tuberculosis-endemic regions. *Trends Mol Med* **18**, 607-614.

Grana, M., Bellinzoni, M., Bellalou, J., Haouz, A., Miras, I., Buschiazzo, A., Winter, N. & Alzari, P. M. (2010). Crystal structure of Mycobacterium tuberculosis LppA, a lipoprotein confined to pathogenic mycobacteria. *Proteins* **78**, 769-772.

Gregory, M. A., Till, R. & Smith, M. C. (2003). Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors. *J Bacteriol* **185**, 5320-5323.

Grundling, A. & Schneewind, O. (2007). Synthesis of glycerol phosphate lipoteichoic acid in Staphylococcus aureus. *Proc Natl Acad Sci U S A* **104**, 8478-8483.

Gupta, R., Lavollay, M., Mainardi, J. L., Arthur, M., Bishai, W. R. & Lamichhane, G. (2010). The Mycobacterium tuberculosis protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. *Nature medicine* **16**, 466-469.

Gupta, R. S. (1998). What are archaebacteria: life's third domain or monoderm prokaryotes related to gram-positive bacteria? A new proposal for the classification of prokaryotic organisms. *Mol Microbiol* **29**, 695-707.

Gupta, S. D., Dowhan, W. & Wu, H. C. (1991). Phosphatidylethanolamine is not essential for the N-acylation of apolipoprotein in Escherichia coli. *J Biol Chem* **266**, 9983-9986.

Gupta, U. D., Katoch, V. M. & McMurray, D. N. (2007). Current status of TB vaccines. *Vaccine* **25**, 3742-3751.

Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. (2003).PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. In Proc Natl Acad Sci USA, pp. 1541-1546.

Haake, D. A. (2000). Spirochaetal lipoproteins and pathogenesis. *Microbiology* **146 (Pt 7)**, 1491-1504.

Hagan, C. L., Kim, S. & Kahne, D. (2010). Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890-892.

Hagan, C. L. & Kahne, D. (2011). The reconstituted Escherichia coli Bam complex catalyzes multiple rounds of beta-barrel assembly. *Biochemistry* **50**, 7444-7446.

Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). beta-Barrel membrane protein assembly by the Bam complex. *Annual review of biochemistry* **80**, 189-210.

Hara, T., Matsuyama, S. & Tokuda, H. (2003). Mechanism underlying the inner membrane retention of Escherichia coli lipoproteins caused by Lol avoidance signals. *J Biol Chem* **278**, 40408-40414.

Harding, C. V. & Boom, W. H. (2010). Regulation of antigen presentation by Mycobacterium tuberculosis: a role for Toll-like receptors. *Nat Rev Microbiol* **8**, 296-307.

Hashimoto, M., Asai, Y. & Ogawa, T. (2004). Separation and structural analysis of lipoprotein in a lipopolysaccharide preparation from Porphyromonas gingivalis. *International immunology* **16**, 1431-1437.

Hashimoto, M., Tawaratsumida, K., Kariya, H., Kiyohara, A., Suda, Y., Krikae, F., Kirikae, T. & Gotz, F. (2006). Not lipoteichoic acid but lipoproteins appear to be the dominant immunobiologically active compounds in Staphylococcus aureus. *J Immunol* **177**, 3162-3169.

Hayashi, S., Chang, S. Y., Chang, S., Giam, C. Z. & Wu, H. C. (1985). Modification and processing of internalized signal sequences of prolipoprotein in Escherichia coli and in Bacillus subtilis. *J Biol Chem* **260**, 5753-5759.

Henao-Tamayo, M., Junqueira-Kipnis, A. P., Ordway, D., Gonzales-Juarrero, M., Stewart, G. R., Young, D. B., Wilkinson, R. J., Basaraba, R. J. & Orme, I. M. (2007). A mutant of Mycobacterium tuberculosis lacking the 19-kDa lipoprotein Rv3763 is highly attenuated in vivo but retains potent vaccinogenic properties. *Vaccine* **25**, 7153-7159.

Henneke, P., Dramsi, S., Mancuso, G. & other authors (2008). Lipoproteins are critical TLR2 activating toxins in group B streptococcal sepsis. *J Immunol* **180**, 6149-6158.

Herrmann, J. L., O'Gaora, P., Gallagher, A., Thole, J. E. & Young, D. B. (1996). Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from Mycobacterium tuberculosis. *Embo J* **15**, 3547-3554.

Hillmann, F., Argentini, M. & Buddelmeijer, N. (2011). Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. *J Biol Chem* **286**, 27936-27946.

Hirano, S., Asamizu, S., Onaka, H., Shiro, Y. & Nagano, S. (2008). Crystal structure of VioE, a key player in the construction of the molecular skeleton of violacein. *J Biol Chem* **283**, 6459-6466.

Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M. & Engelhardt, H. (2008). Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* **105**, 3963-3967.

Hoshino, K., Takeuchi, O., Kawai, T., Sanjo, H., Ogawa, T., Takeda, Y., Takeda, K. & Akira, S. (1999). Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J Immunol* **162**, 3749-3752.

Hoskisson, P. A. & Hutchings, M. I. (2006).MtrAB-LpqB: a conserved three-component system in actinobacteria? In Trends Microbiol, pp. 444-449.

Hovav, A. H., Mullerad, J., Davidovitch, L., Fishman, Y., Bigi, F., Cataldi, A. & Bercovier, H. (2003). The Mycobacterium tuberculosis recombinant 27-kilodalton lipoprotein induces a strong Th1-type immune response deleterious to protection. *Infect Immun* **71**, 3146-3154.

Huguet-Tapia, J. C., Badger, J. H., Loria, R. & Pettis, G. S. (2011). Streptomyces turgidiscabies Car8 contains a modular pathogenicity island that shares virulence genes with other actinobacterial plant pathogens. *Plasmid* **65**, 118-124.

Hutchings, M. I. (2007). Unusual two-component signal transduction pathways in the actinobacteria. *Adv Appl Microbiol* **61**, 1-26.

Hutchings, M. I., Palmer, T., Harrington, D. J. & Sutcliffe, I. C. (2009).Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. In Trends Microbiol, pp. 13-21.

Inukai, M., Nakajima, M., Osawa, M., Haneishi, T. & Arai, M. (1978). Globomycin, a new peptide antibiotic with spheroplast-forming activity. II. Isolation and physico-chemical and biological characterization. *The Journal of antibiotics* **31**, 421-425.

Ip, W. K., Sokolovska, A., Charriere, G. M. & other authors (2010). Phagocytosis and phagosome acidification are required for pathogen processing and MyD88-dependent responses to Staphylococcus aureus. *J Immunol* **184**, 7071-7081.

Ito, Y., Kanamaru, K., Taniguchi, N., Miyamoto, S. & Tokuda, H. (2006). A novel ligand bound ABC transporter, LolCDE, provides insights into the molecular mechanisms underlying membrane detachment of bacterial lipoproteins. *Mol Microbiol* **62**, 1064-1075.

Iwasaki, A. & Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. *Science* **327**, 291-295.

Jackowski, S. & Rock, C. O. (1986). Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of Escherichia coli. *J Biol Chem* **261**, 11328-11333.

Jarlier, V. & Nikaido, H. (1990). Permeability barrier to hydrophilic solutes in Mycobacterium chelonei. *J Bacteriol* **172**, 1418-1423.

Jin, M. S., Kim, S. E., Heo, J. Y., Lee, M. E., Kim, H. M., Paik, S. G., Lee, H. & Lee, J. O. (2007). Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**, 1071-1082.

Joshi, M. V., Bignell, D. R., Johnson, E. G., Sparks, J. P., Gibson, D. M. & Loria, R. (2007). The AraC/XylS regulator TxtR modulates thaxtomin biosynthesis and virulence in Streptomyces scabies. *Mol Microbiol* **66**, 633-642.

Joshi, M. V., Mann, S. G., Antelmann, H. & other authors (2010). The twin arginine protein transport pathway exports multiple virulence proteins in the plant pathogen Streptomyces scabies. *Mol Microbiol* **77**, 252-271.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein science : a publication of the Protein Society* **12**, 1652-1662.

Kalscheuer, R., Weinrick, B., Veeraraghavan, U., Besra, G. S. & Jacobs, W. R., Jr. (2010). Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **107**, 21761-21766.

Kang, H. J., Ha, J. M., Kim, H. S., Lee, H., Kurokawa, K. & Lee, B. L. (2011). The role of phagocytosis in IL-8 production by human monocytes in response to lipoproteins on Staphylococcus aureus. *Biochem Biophys Res Commun* **406**, 449-453.

Kang, J. Y., Nan, X., Jin, M. S. & other authors (2009). Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **31**, 873-884.

Karakousis, P. C., Bishai, W. R. & Dorman, S. E. (2004). Mycobacterium tuberculosis cell envelope lipids and the host immune response. *Cellular microbiology* **6**, 105-116.

Kaufmann, S. H., Hussey, G. & Lambert, P. H. (2010). New vaccines for tuberculosis. *Lancet* **375**, 2110-2119.

Kaufmann, S. H. & Gengenbacher, M. (2012). Recombinant live vaccine candidates against tuberculosis. *Curr Opin Biotechnol* **23**, 900-907.

Kieser, T., Bibb, MJ, Buttner, MJ, Chater, KF, Hopwood, DA (2000). Practical Streptomyces Genetics.

Kiho, T., Nakayama, M., Yasuda, K., Miyakoshi, S., Inukai, M. & Kogen, H. (2003). Synthesis and antimicrobial activity of novel globomycin analogues. *Bioorganic & medicinal chemistry letters* **13**, 2315-2318.

Kiho, T., Nakayama, M., Yasuda, K., Miyakoshi, S., Inukai, M. & Kogen, H. (2004). Structure-activity relationships of globomycin analogues as antibiotics. *Bioorganic & medicinal chemistry* **12**, 337-361.

Kim, D. Y., Jin, K. S., Kwon, E., Ree, M. & Kim, K. K. (2007). Crystal structure of RseB and a model of its binding mode to RseA. *Proc Natl Acad Sci U S A* **104**, 8779-8784.

Klein, J. R., Henrich, B. & Plapp, R. (1991). Molecular analysis and nucleotide sequence of the envCD operon of Escherichia coli. *Molecular & general genetics : MGG* **230**, 230-240.

Kobayashi, N., Nishino, K., Hirata, T. & Yamaguchi, A. (2003). Membrane topology of ABC-type macrolide antibiotic exporter MacB in Escherichia coli. *FEBS Lett* **546**, 241-246.

Kovacevic, S., Anderson, D., Morita, Y. S., Patterson, J., Haites, R., McMillan, B. N., Coppel, R., McConville, M. J. & Billman-Jacobe, H. (2006). Identification of a novel protein with a role in lipoarabinomannan biosynthesis in mycobacteria. *J Biol Chem* **281**, 9011-9017.

Kovacs-Simon, A., Titball, R. W. & Michell, S. L. (2011). Lipoproteins of bacterial pathogens. *Infect Immun* **79**, 548-561.

Kurokawa, K., Lee, H., Roh, K. B. & other authors (2009). The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of Staphylococcus aureus Functions as a Native Ligand for Toll-like Receptor 2. *J Biol Chem* **284**, 8406-8411.

Kurokawa, K., Kim, M. S., Ichikawa, R., Ryu, K. H., Dohmae, N., Nakayama, H. & Lee, B. L. (2012a). Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in Staphylococcus aureus. *J Bacteriol* **194**, 3299-3306.

Kurokawa, K., Ryu, K. H., Ichikawa, R. & other authors (2012b). Novel bacterial lipoprotein structures conserved in low-GC content gram-positive bacteria are recognized by Toll-like receptor 2. *J Biol Chem* **287**, 13170-13181.

Lai, J. S. & Wu, H. C. (1980). Incorporation of acyl moieties of phospholipids into murein lipoprotein in intact cells of Escherichia coli by phospholipid vesicle fusion. *J Bacteriol* **144**, 451-453.

Lawn, S. D. & Zumla, A. I. (2011). Tuberculosis. *Lancet* **378**, 57-72.

Lewenza, S., Vidal-Ingigliardi, D. & Pugsley, A. P. (2006). Direct visualization of red fluorescent lipoproteins indicates conservation of the membrane sorting rules in the family Enterobacteriaceae. *J Bacteriol* **188**, 3516-3524.

Li, Y., Powell, D. A., Shaffer, S. A. & other authors (2012). LPS remodeling is an evolved survival strategy for bacteria. *Proc Natl Acad Sci U S A* **109**, 8716-8721.

Lin, J. J., Kanazawa, H. & Wu, H. C. (1980). Assembly of outer membrane lipoprotein in an Escherichia coli mutant with a single amino acid replacement within the signal sequence of prolipoprotein. *J Bacteriol* **141**, 550-557.

Liu, C. F., Tonini, L., Malaga, W. & other authors (2013). Bacterial protein-O-mannosylating enzyme is crucial for virulence of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **110**, 6560-6565.

Lowy, F. D. (1998). Staphylococcus aureus infections. *The New England journal of medicine* **339**, 520-532.

Machata, S., Tchatalbachev, S., Mohamed, W., Jansch, L., Hain, T. & Chakraborty, T. (2008). Lipoproteins of Listeria monocytogenes are critical for virulence and TLR2-mediated immune activation. *J Immunol* **181**, 2028-2035.

Madigan, M. T., Martinko, J.M., Parker, J. (2000). In Brock- Biology of Microorganisms, pp. 49-67. Edited by I. C.P.F.: Prentice-Hall.

Mahapatra, S., Basu, J., Brennan, P.J., and Crick, D.C. (2005). Structure, biosynthesis, and genetics of the mycolic acid-arabinogalactan-peptidoglycan complex. In Tuberculosis and the Tubercle Bacillus, pp. 275-285. Edited by S. T. Cole, Eisenach, K.D., MaMurray, D.N., and Jacobs, W.R. Washington, D.C.: ASM Press.

Makarov, V., Manina, G., Mikusova, K. & other authors (2009). Benzothiazinones kill Mycobacterium tuberculosis by blocking arabinan synthesis. *Science* **324**, 801-804.

Malen, H., Pathak, S., Softeland, T., de Souza, G. A. & Wiker, H. G. (2010). Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv. *BMC microbiology* **10**, 132.

Malinverni, J. C. & Silhavy, T. J. (2009). An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc Natl Acad Sci U S A* **106**, 8009-8014.

Marland, Z., Beddoe, T., Zaker-Tabrizi, L. & other authors (2006). Hijacking of a substrate-binding protein scaffold for use in mycobacterial cell wall biosynthesis. *J Mol Biol* **359**, 983-997.

Master, S. S., Rampini, S. K., Davis, A. S., Keller, C., Ehlers, S., Springer, B., Timmins, G. S., Sander, P. & Deretic, V. (2008). Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**, 224-232.

Masuda, K., Matsuyama, S. & Tokuda, H. (2002). Elucidation of the function of lipoprotein-sorting signals that determine membrane localization. *Proc Natl Acad Sci U S A* **99**, 7390-7395.

Matsuyama, S., Tajima, T. & Tokuda, H. (1995). A novel periplasmic carrier protein involved in the sorting and transport of Escherichia coli lipoproteins destined for the outer membrane. *EMBO J* **14**, 3365-3372.

Matsuyama, S., Yokota, N. & Tokuda, H. (1997). A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli. *EMBO J* **16**, 6947-6955.

Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., Bradbury, E. M., Bradbury, A. R. & Chen, X. (2005). Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Mol Biol Cell* **16**, 396-404.

Mayerle, J., den Hoed, C. M., Schurmann, C. & other authors (2013). Identification of genetic loci associated with Helicobacter pylori serologic status. *Jama* **309**, 1912-1920.

McDonough, J. A., Hacker, K. E., Flores, A. R., Pavelka, M. S., Jr. & Braunstein, M. (2005). The twin-arginine translocation pathway of Mycobacterium smegmatis is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol* **187**, 7667-7679.

Migliori, G. B., De Iaco, G., Besozzi, G., Centis, R. & Cirillo, D. M. (2007). First tuberculosis cases in Italy resistant to all tested drugs. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **12**, E070517 070511.

Minnikin, D. E. (1982). Lipids: complex lipids, their chemistry, biosynthesis and roles. In The biology of the Mycobacteria, pp. 95-184. Edited by C. Ratledge, and Stanford, J., . New York: Academic Press. Inc.

Miyadai, H., Tanaka-Masuda, K., Matsuyama, S. & Tokuda, H. (2004). Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the Escherichia coli periplasm. *J Biol Chem* **279**, 39807-39813.

Miyamoto, A., Matsuyama, S. & Tokuda, H. (2001). Mutant of LolA, a lipoprotein-specific molecular chaperone of Escherichia coli, defective in the transfer of lipoproteins to LolB. *Biochem Biophys Res Commun* **287**, 1125-1128.

Miyamoto, S. & Tokuda, H. (2007). Diverse effects of phospholipids on lipoprotein sorting and ATP hydrolysis by the ABC transporter LolCDE complex. *Biochimica et biophysica acta* **1768**, 1848-1854.

Mohiman, N., Argentini, M., Batt, S. M., Cornu, D., Masi, M., Eggeling, L., Besra, G. & Bayan, N. (2012). The ppm operon is essential for acylation and glycosylation of lipoproteins in Corynebacterium glutamicum. *PLoS One* **7**, e46225.

Muhlradt, P. F., Kiess, M., Meyer, H., Sussmuth, R. & Jung, G. (1997). Isolation, structure elucidation, and synthesis of a macrophage stimulatory lipopeptide from Mycoplasma fermentans acting at picomolar concentration. *The Journal of experimental medicine* **185**, 1951-1958.

Mukhopadhyay, S. & Balaji, K. N. (2011). The PE and PPE proteins of Mycobacterium tuberculosis. *Tuberculosis (Edinb)* **91**, 441-447.

Murphy, R. C. & Gaskell, S. J. (2011). New applications of mass spectrometry in lipid analysis. *J Biol Chem* **286**, 25427-25433.

Nakada, S., Sakakura, M., Takahashi, H., Okuda, S., Tokuda, H. & Shimada, I. (2009). Structural investigation of the interaction between LolA and LolB using NMR. *J Biol Chem* **284**, 24634-24643.

Nakayama, H., Kurokawa, K. & Lee, B. L. (2012). Lipoproteins in bacteria: structures and biosynthetic pathways. *Febs J* **279**, 4247-4268.

Narita, S., Tanaka, K., Matsuyama, S. & Tokuda, H. (2002). Disruption of lolCDE, encoding an ATP-binding cassette transporter, is lethal for Escherichia coli and prevents release of lipoproteins from the inner membrane. *J Bacteriol* **184**, 1417-1422.

Narita, S., Kanamaru, K., Matsuyama, S. & Tokuda, H. (2003). A mutation in the membrane subunit of an ABC transporter LolCDE complex causing outer membrane localization of lipoproteins against their inner membrane-specific signals. *Mol Microbiol* **49**, 167-177.

Narita, S. & Tokuda, H. (2007). Amino acids at positions 3 and 4 determine the membrane specificity of Pseudomonas aeruginosa lipoproteins. *J Biol Chem* **282**, 13372-13378.

Narita, S. & Tokuda, H. (2011). Overexpression of LolCDE allows deletion of the Escherichia coli gene encoding apolipoprotein N-acyltransferase. *J Bacteriol* **193**, 4832-4840.

Navarre, W. W., Daefler, S. & Schneewind, O. (1996). Cell wall sorting of lipoproteins in Staphylococcus aureus. *J Bacteriol* **178**, 441-446.

Nguyen, H. T., Wolff, K. A., Cartabuke, R. H., Ogwang, S. & Nguyen, L. (2010). A lipoprotein modulates activity of the MtrAB two-component system to provide intrinsic multidrug resistance, cytokinetic control and cell wall homeostasis in Mycobacterium. *Mol Microbiol* **76**, 348-364.

Nguyen, K. T., Piastro, K. & Derbyshire, K. M. (2009). LpqM, a mycobacterial lipoprotein-metalloproteinase, is required for conjugal DNA transfer in Mycobacterium smegmatis. *J Bacteriol* **191**, 2721-2727.

Niederweis, M., Danilchanka, O., Huff, J., Hoffmann, C. & Engelhardt, H. (2010). Mycobacterial outer membranes: in search of proteins. *Trends Microbiol* **18**, 109-116.

Nielsen, J. B. & Lampen, J. O. (1983). Beta-lactamase III of Bacillus cereus 569: membrane lipoprotein and secreted protein. *Biochemistry* **22**, 4652-4656.

Nikaido, H. (1998). Multiple antibiotic resistance and efflux. *Current opinion in microbiology* **1**, 516-523.

Noss, E. H., Pai, R. K., Sellati, T. J., Radolf, J. D., Belisle, J., Golenbock, D. T., Boom, W. H. & Harding, C. V. (2001). Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kDa lipoprotein of Mycobacterium tuberculosis. *J Immunol* **167**, 910-918.

O'Garra, A., Redford, P. S., McNab, F. W., Bloom, C. I., Wilkinson, R. J. & Berry, M. P. (2013). The immune response in tuberculosis. *Annual review of immunology* **31**, 475-527.

Oguchi, Y., Takeda, K., Watanabe, S., Yokota, N., Miki, K. & Tokuda, H. (2008). Opening and closing of the hydrophobic cavity of LolA coupled to lipoprotein binding and release. *J Biol Chem* **283**, 25414-25420.

Oku, Y., Kurokawa, K., Matsuo, M., Yamada, S., Lee, B. L. & Sekimizu, K. (2009). Pleiotropic roles of polyglycerolphosphate synthase of lipoteichoic acid in growth of Staphylococcus aureus cells. *J Bacteriol* **191**, 141-151.

Okuda, S., Watanabe, S. & Tokuda, H. (2008). A short helix in the C-terminal region of LolA is important for the specific membrane localization of lipoproteins. *FEBS Lett* **582**, 2247-2251.

Okuda, S. & Tokuda, H. (2009). Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc Natl Acad Sci U S A* **106**, 5877-5882.

Okuda, S. & Tokuda, H. (2011). Lipoprotein sorting in bacteria. *Annu Rev Microbiol* **65**, 239-259.

Okuyama, H., Kankura, T. & Nojima, S. (1967). Positional distribution of fatty acids in phospholipids from Mycobacteria. *J Biochem* **61**, 732-737.

Omueti, K. O., Beyer, J. M., Johnson, C. M., Lyle, E. A. & Tapping, R. I. (2005). Domain exchange between human toll-like receptors 1 and 6 reveals a region required for lipopeptide discrimination. *J Biol Chem* **280**, 36616-36625.

Paradis-Bleau, C., Markovski, M., Uehara, T., Lupoli, T. J., Walker, S., Kahne, D. E. & Bernhardt, T. G. (2010). Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. *Cell* **143**, 1110-1120.

Pathania, R., Zlitni, S., Barker, C. & other authors (2009). Chemical genomics in Escherichia coli identifies an inhibitor of bacterial lipoprotein targeting. *Nature chemical biology* **5**, 849-856.

Prados-Rosales, R., Baena, A., Martinez, L. R. & other authors (2011). Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. *J Clin Invest* **121**, 1471-1483.

Prados-Rosales, R., Baena, A., Martinez, L. R. & other authors (2013). Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. *J Clin Invest* **121**, 1471-1483.

Raivio, T. L. & Silhavy, T. J. (2001). Periplasmic stress and ECF sigma factors. *Annu Rev Microbiol* **55**, 591-624.

Ramesh, G., Alvarez, A. L., Roberts, E. D., Dennis, V. A., Lasater, B. L., Alvarez, X. & Philipp, M. T. (2003). Pathogenesis of Lyme neuroborreliosis: Borrelia burgdorferi lipoproteins induce both proliferation and apoptosis in rhesus monkey astrocytes. *European journal of immunology* **33**, 2539-2550.

Ramos, H. C., Rumbo, M. & Sirard, J. C. (2004). Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends Microbiol* **12**, 509-517.

Rampini, S. K., Selchow, P., Keller, C., Ehlers, S., Bottger, E. C. & Sander, P. (2008). LspA inactivation in Mycobacterium tuberculosis results in attenuation without affecting phagosome maturation arrest. *Microbiology* **154**, 2991-3001.

Ray, A., Cot, M., Puzo, G., Gilleron, M. & Nigou, J. (2013). Bacterial cell wall macroamphiphiles: pathogen-/microbe-associated molecular patterns detected by mammalian innate immune system. *Biochimie* **95**, 33-42.

Remans, K., Pauwels, K., van Ulsen, P., Buts, L., Cornelis, P., Tommassen, J., Savvides, S. N., Decanniere, K. & Van Gelder, P. (2010). Hydrophobic surface patches on LolA of Pseudomonas aeruginosa are essential for lipoprotein binding. *J Mol Biol* **401**, 921-930.

Rezwan, M., Grau, T., Tschumi, A. & Sander, P. (2007a). Lipoprotein synthesis in mycobacteria. *Microbiology* **153**, 652-658.

Rezwan, M., Laneelle, M. A., Sander, P. & Daffe, M. (2007b). Breaking down the wall: fractionation of mycobacteria. *Journal of microbiological methods* **68**, 32-39.

Rietveld, A. G., Koorengevel, M. C. & de Kruijff, B. (1995). Non-bilayer lipids are required for efficient protein transport across the plasma membrane of Escherichia coli. *EMBO J* **14**, 5506-5513.

Robichon, C., Vidal-Ingigliardi, D. & Pugsley, A. P. (2005). Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in Escherichia coli. *J Biol Chem* **280**, 974-983.

Ruiz, N., Gronenberg, L. S., Kahne, D. & Silhavy, T. J. (2008). Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *Proc Natl Acad Sci U S A* **105**, 5537-5542.

Russell, D. G. (2001). Mycobacterium tuberculosis: here today, and here tomorrow. *Nat Rev Mol Cell Biol* **2**, 569-577.

Russell, D. G. (2007). Who puts the tubercle in tuberculosis? *Nat Rev Microbiol* **5**, 39-47.

Russell, D. G. (2011). Mycobacterium tuberculosis and the intimate discourse of a chronic infection. *Immunological reviews* **240**, 252-268.

Ryan, K. S., Balibar, C. J., Turo, K. E., Walsh, C. T. & Drennan, C. L. (2008). The violacein biosynthetic enzyme VioE shares a fold with lipoprotein transporter proteins. *J Biol Chem* **283**, 6467-6475.

Ryu, Y. & Schultz, P. G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nature methods* **3**, 263-265.

Sakamoto, C., Satou, R., Tokuda, H. & Narita, S. (2010). Novel mutations of the LolCDE complex causing outer membrane localization of lipoproteins despite their inner membrane-retention signals. *Biochem Biophys Res Commun* **401**, 586-591.

Sakthi, S. & Narayanan, S. (2013). The lpqS knockout mutant of Mycobacterium tuberculosis is attenuated in Macrophages. *Microbiol Res*.

Sander, P., Meier, A. & Bottger, E. C. (1995). rpsL+: a dominant selectable marker for gene replacement in mycobacteria. *Mol Microbiol* **16**, 991-1000.

Sander, P., Prammananan, T., Meier, A., Frischkorn, K. & Bottger, E. C. (1997). The role of ribosomal RNAs in macrolide resistance. *Mol Microbiol* **26**, 469-480.

Sander, P., Springer, B. & Bottger, E. C. (2001). Gene Replacement in Mycobacterium tuberculosis and Mycobacterium bovis BCG Using rpsL as a Dominant Negative Selectable Marker. *Methods Mol Med* **54**, 93-104.

Sander, P., Rezwan, M., Walker, B. & other authors (2004). Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. *Mol Microbiol* **52**, 1543-1552.

Sankaran, K. & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J Biol Chem* **269**, 19701-19706.

Sartain, M. J. & Belisle, J. T. (2009). N-Terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. *Glycobiology* **19**, 38-51.

Sartain, M. J., Dick, D. L., Rithner, C. D., Crick, D. C. & Belisle, J. T. (2011). Lipidomic analyses of Mycobacterium tuberculosis based on accurate mass measurements and the novel "Mtb LipidDB". *Journal of lipid research* **52**, 861-872.

Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84.

Sassetti, C. M. & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**, 12989-12994.

Schlosser, A., Kampers, T. & Schrempf, H. (1997). The Streptomyces ATP-binding component MsiK assists in cellobiose and maltose transport. *J Bacteriol* **179**, 2092-2095.

Schmaler, M., Jann, N. J., Gotz, F. & Landmann, R. (2010). Staphylococcal lipoproteins and their role in bacterial survival in mice. *International journal of medical microbiology : IJMM* **300**, 155-160.

Schulze, R. J., Chen, S., Kumru, O. S. & Zuckert, W. R. (2010). Translocation of Borrelia burgdorferi surface lipoprotein OspA through the outer membrane requires an unfolded conformation and can initiate at the C-terminus. *Mol Microbiol* **76**, 1266-1278.

Scragg, I. G., Kwiatkowski, D., Vidal, V., Reason, A., Paxton, T., Panico, M., Dell, A. & Morris, H. (2000). Structural characterization of the inflammatory moiety of a variable major lipoprotein of Borrelia recurrentis. *J Biol Chem* **275**, 937-941.

Seiffer, D., Klein, J. R. & Plapp, R. (1993). EnvC, a new lipoprotein of the cytoplasmic membrane of Escherichia coli. *FEMS microbiology letters* **107**, 175-178.

Seltmann, G., Holst, O. (2002). In The Bacterial Cell Wall, pp. 103-161: Springer.

Serebryakova, M. V., Demina, I. A., Galyamina, M. A., Kondratov, I. G., Ladygina, V. G. & Govorun, V. M. (2011). The acylation state of surface lipoproteins of mollicute Acholeplasma laidlawii. *J Biol Chem* **286**, 22769-22776.

Shimada, T., Park, B. G., Wolf, A. J. & other authors (2010). Staphylococcus aureus evades lysozyme-based peptidoglycan digestion that links phagocytosis, inflammasome activation, and IL-1beta secretion. *Cell Host Microbe* **7**, 38-49.

Shimizu, T., Kida, Y. & Kuwano, K. (2007). Triacylated lipoproteins derived from Mycoplasma pneumoniae activate nuclear factor-kappaB through toll-like receptors 1 and 2. *Immunology* **121**, 473-483.

Shimizu, T., Kida, Y. & Kuwano, K. (2008). A triacylated lipoprotein from Mycoplasma genitalium activates NF-kappaB through Toll-like receptor 1 (TLR1) and TLR2. *Infect Immun* **76**, 3672-3678.

Shruthi, H., Babu, M. M. & Sankaran, K. (2010). TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. *Journal of molecular evolution* **70**, 359-370.

Simeone, R., Bobard, A., Lippmann, J., Bitter, W., Majlessi, L., Brosch, R. & Enninga, J. (2012). Phagosomal rupture by Mycobacterium tuberculosis results in toxicity and host cell death. *PLoS Pathog* **8**, e1002507.

Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. *Proc Natl Acad Sci U S A* **104**, 6400-6405.

Smith, P. C., Karpowich, N., Millen, L., Moody, J. E., Rosen, J., Thomas, P. J. & Hunt, J. F. (2002). ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Molecular cell* **10**, 139-149.

Sperandeo, P., Cescutti, R., Villa, R., Di Benedetto, C., Candia, D., Deho, G. & Polissi, A. (2007). Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli. *J Bacteriol* **189**, 244-253.

Sperandeo, P., Deho, G. & Polissi, A. (2009). The lipopolysaccharide transport system of Gram-negative bacteria. *Biochimica et biophysica acta* **1791**, 594-602.

Steyn, A. J., Joseph, J. & Bloom, B. R. (2003). Interaction of the sensor module of Mycobacterium tuberculosis H37Rv KdpD with members of the Lpr family. *Mol Microbiol* **47**, 1075-1089.

Stoll, H., Dengjel, J., Nerz, C. & Gotz, F. (2005). Staphylococcus aureus deficient in lipidation of prelipoproteins is attenuated in growth and immune activation. *Infect Immun* **73**, 2411-2423.

Stover, C. K., de la Cruz, V. F., Fuerst, T. R. & other authors (1991). New use of BCG for recombinant vaccines. *Nature* **351**, 456-460.

Sulzenbacher, G., Canaan, S., Bordat, Y. & other authors (2006). LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of Mycobacterium tuberculosis. *Embo J* **25**, 1436-1444.

Sutcliffe, I. C. & Russell, R. R. (1995). Lipoproteins of gram-positive bacteria. *J Bacteriol* **177**, 1123-1128.

Sutcliffe, I. C. & Harrington, D. J. (2002). Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology* **148**, 2065-2077.

Sutcliffe, I. C. & Harrington, D. J. (2004). Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* **28**, 645-659.

Sutcliffe, I. C. & Hutchings, M. I. (2007). Putative lipoproteins identified by bioinformatic genome analysis of Leifsonia xyli ssp. xyli, the causative agent of sugarcane ratoon stunting disease. *Mol Plant Pathol* **8**, 121-128.

Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol* **18**, 464-470.

Tajima, T., Yokota, N., Matsuyama, S. & Tokuda, H. (1998). Genetic analyses of the in vivo function of LolA, a periplasmic chaperone involved in the outer membrane localization of Escherichia coli lipoproteins. *FEBS Lett* **439**, 51-54.

Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S., Tokuda, H. & Miki, K. (2003). Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. *EMBO J* **22**, 3199-3209.

Takeuchi, O., Hoshino, K., Kawai, T., Sanjo, H., Takada, H., Ogawa, T., Takeda, K. & Akira, S. (1999). Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity* **11**, 443-451.

Takeuchi, O., Kawai, T., Muhlradt, P. F., Morr, M., Radolf, J. D., Zychlinsky, A., Takeda, K. & Akira, S. (2001). Discrimination of bacterial lipoproteins by Toll-like receptor 6. *International immunology* **13**, 933-940.

Takeuchi, O., Sato, S., Horiuchi, T., Hoshino, K., Takeda, K., Dong, Z., Modlin, R. L. & Akira, S. (2002). Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J Immunol* **169**, 10-14.

Takeuchi, O. & Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell* **140**, 805-820.

Tanaka, K., Matsuyama, S. I. & Tokuda, H. (2001). Deletion of lolB, encoding an outer membrane lipoprotein, is lethal for Escherichia coli and causes accumulation of lipoprotein localization intermediates in the periplasm. *J Bacteriol* **183**, 6538-6542.

Tanaka, S. Y., Narita, S. & Tokuda, H. (2007). Characterization of the Pseudomonas aeruginosa Lol system as a lipoprotein sorting mechanism. *J Biol Chem* **282**, 13379-13384.

Taniguchi, N., Matsuyama, S. & Tokuda, H. (2005). Mechanisms underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed {beta}-barrel structures. *J Biol Chem* **280**, 34481-34488.

Tao, K., Watanabe, S., Narita, S. & Tokuda, H. (2010). A periplasmic LolA derivative with a lethal disulfide bond activates the Cpx stress response system. *J Bacteriol* **192**, 5657-5662.

Tawaratsumida, K., Furuyashiki, M., Katsumoto, M., Fujimoto, Y., Fukase, K., Suda, Y. & Hashimoto, M. (2009). Characterization of N-terminal structure of TLR2-activating lipoprotein in Staphylococcus aureus. *J Biol Chem* **284**, 9147-9152.

Terada, M., Kuroda, T., Matsuyama, S. I. & Tokuda, H. (2001). Lipoprotein sorting signals evaluated as the LolA-dependent release of lipoproteins from the cytoplasmic membrane of Escherichia coli. *J Biol Chem* **276**, 47690-47694.

Thompson, B. J., Widdick, D. A., Hicks, M. G., Chandra, G., Sutcliffe, I. C., Palmer, T. & Hutchings, M. I. (2010). Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium Streptomyces coelicolor. *Mol Microbiol* **77**, 943-957.

Tjalsma, H. & van Dijl, J. M. (2005). Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics* **5**, 4472-4482.

Tokuda, H. & Matsuyama, S. (2004). Sorting of lipoproteins to the outer membrane in E. coli. *Biochimica et biophysica acta* **1693**, 5-13.

Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Bioscience, biotechnology, and biochemistry* **73**, 465-473.

Traag, B. A., Driks, A., Stragier, P. & other authors (2010). Do mycobacteria produce endospores? *Proc Natl Acad Sci U S A* **107**, 878-881.

Tschumi, A., Nai, C., Auchli, Y., Hunziker, P., Gehrig, P., Keller, P., Grau, T. & Sander, P. (2009). Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria. *J Biol Chem* **284**, 27146-27156.

Tschumi, A., Grau, T., Albrecht, D., Rezwan, M., Antelmann, H. & Sander, P. (2012). Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial lgt mutant. *J Bacteriol* **194**, 3938-3949.

Tsukahara, J., Mukaiyama, K., Okuda, S., Narita, S. & Tokuda, H. (2009). Dissection of LolB function--lipoprotein binding, membrane targeting and incorporation of lipoproteins into lipid bilayers. *FEBS J* **276**, 4496-4504.

Typas, A., Banzhaf, M., van den Berg van Saparoea, B. & other authors (2010). Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* **143**, 1097-1109.

Ujihara, T., Sakurai, I., Mizusawa, N. & Wada, H. (2008). A method for analyzing lipid-modified proteins with mass spectrometry. *Anal Biochem* **374**, 429-431.

van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., Brenner, M. & Peters, P. J. (2007). M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**, 1287-1298.

Vandal, O. H., Pierini, L. M., Schnappinger, D., Nathan, C. F. & Ehrt, S. (2008). A membrane protein preserves intrabacterial pH in intraphagosomal Mycobacterium tuberculosis. *Nature medicine* **14**, 849-854.

Velayati, A. A., Farnia, P., Masjedi, M. R. & other authors (2009). Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. *The European respiratory journal* **34**, 1202-1203.

Vidal-Ingigliardi, D., Lewenza, S. & Buddelmeijer, N. (2007). Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. *J Bacteriol* **189**, 4456-4464.

Volokhina, E. B., Beckers, F., Tommassen, J. & Bos, M. P. (2009). The beta-barrel outer membrane protein assembly complex of Neisseria meningitidis. *J Bacteriol* **191**, 7074-7085.

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262-265.

Voulhoux, R. & Tommassen, J. (2004). Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Research in microbiology* **155**, 129-135.

Wang, L., Xie, J. & Schultz, P. G. (2006). Expanding the genetic code. *Annual review of biophysics and biomolecular structure* **35**, 225-249.

Watanabe, S., Oguchi, Y., Yokota, N. & Tokuda, H. (2007). Large-scale preparation of the homogeneous LolA lipoprotein complex and efficient in vitro transfer of lipoproteins to the outer membrane in a LolB-dependent manner. *Protein science : a publication of the Protein Society* **16**, 2741-2749.

Watanabe, S., Oguchi, Y., Takeda, K., Miki, K. & Tokuda, H. (2008). Introduction of a lethal redox switch that controls the opening and closing of the hydrophobic cavity in LolA. *J Biol Chem* **283**, 25421-25427.

Weiner, J., 3rd & Kaufmann, S. H. (2014). Recent advances towards tuberculosis control: vaccines and biomarkers. *Journal of internal medicine* **275**, 467-480.

WHO TB-factsheet, W. H. O., Geneva (2009). WHO TB-factsheet, World Health Organization, Geneva. *http://wwwwhoint/tb/publications/2009/tbfactsheet_2009_one_pagepdf*.

Widdick, D. A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschröder, M. & Palmer, T. (2006).The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. In Proc Natl Acad Sci USA, pp. 17927-17932.

Widdick, D. A., Hicks, M. G., Thompson, B. J. & other authors (2011). Dissecting the complete lipoprotein biogenesis pathway in Streptomyces scabies. *Mol Microbiol* **80**, 1395-1412.

Wolfe, L. M., Mahaffey, S. B., Kruh, N. A. & Dobos, K. M. (2010). Proteomic definition of the cell wall of Mycobacterium tuberculosis. *Journal of proteome research* **9**, 5816-5826.

Wu, C. H. (1966). Biosynthesis of lipoproteins. In Escherichia coli and Salmonella typhimurium: cellular and molecular biology, pp. 1005-1014. Edited by F. C. Neidhardt, Curtiss, R. III, Ingraham, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M., and Umbarger, H.E. . Washington, D.C.: ASM Press.

Wu, C. H., Tsai-Wu, J. J., Huang, Y. T., Lin, C. Y., Lioua, G. G. & Lee, F. J. (1998). Identification and subcellular localization of a novel Cu,Zn superoxide dismutase of Mycobacterium tuberculosis. *FEBS Lett* **439**, 192-196.

Wu, H. C. (1996). Biosynthesis of lipoproteins. In Escherichia coli and Salmonella typhimurium: cellular and molecular biology, pp. 1005-1014. Edited by F. C. Neidhardt. Washington, DC: American Society for Microbiology.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121**, 235-245.

Xue, T., Stavropoulos, E., Yang, M. & other authors (2004). RNA encoding the MPT83 antigen induces protective immune responses against Mycobacterium tuberculosis infection. *Infect Immun* **72**, 6324-6329.

Yakushi, T., Yokota, N., Matsuyama, S. & Tokuda, H. (1998). LolA-dependent release of a lipid-modified protein from the inner membrane of Escherichia coli requires nucleoside triphosphate. *J Biol Chem* **273**, 32576-32581.

Yakushi, T., Masuda, K., Narita, S., Matsuyama, S. & Tokuda, H. (2000). A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat Cell Biol* **2**, 212-218.

Yamaguchi, K., Yu, F. & Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell* **53**, 423-432.

Yamanaka, H., Kobayashi, H., Takahashi, E. & Okamoto, K. (2008). MacAB is involved in the secretion of Escherichia coli heat-stable enterotoxin II. *J Bacteriol* **190**, 7693-7698.

Yasuda, M., Iguchi-Yokoyama, A., Matsuyama, S., Tokuda, H. & Narita, S. (2009). Membrane topology and functional importance of the periplasmic region of ABC transporter LolCDE. *Bioscience, biotechnology, and biochemistry* **73**, 2310-2316.

Yuan, Y. R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P. J. & Hunt, J. F. (2001). The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. *J Biol Chem* **276**, 32313-32321.

Zahringer, U., Lindner, B., Inamura, S., Heine, H. & Alexander, C. (2008). TLR2 - promiscuous or specific? A critical re-evaluation of a receptor expressing apparent broad specificity. *Immunobiology* **213**, 205-224.

Zhang, Y. J., Ioerger, T. R., Huttenhower, C., Long, J. E., Sassetti, C. M., Sacchettini, J. C. & Rubin, E. J. (2012). Global assessment of genomic regions required for growth in Mycobacterium tuberculosis. *PLoS Pathog* **8**, e1002946.

Ziegenbalg, A., Prados-Rosales, R., Jenny-Avital, E. R., Kim, R. S., Casadevall, A. & Achkar, J. M. (2013). Immunogenicity of mycobacterial vesicles in humans: Identification of a new tuberculosis antibody biomarker. *Tuberculosis (Edinb)*.

Zuber, B., Haenni, M., Ribeiro, T., Minnig, K., Lopes, F., Moreillon, P. & Dubochet, J. (2006). Granular layer in the periplasmic space of gram-positive bacteria and fine structures of Enterococcus gallinarum and Streptococcus gordonii septa revealed by cryo-electron microscopy of vitreous sections. *J Bacteriol* **188**, 6652-6660.

Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G. & Daffe, M. (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* **190**, 5672-5680.

PUBLICATIONS

2014 Tokuda, H., Sander, P., Lee, B.L., Okuda, S., Grau, T., Tschumi, A., **Brülle, J.K.**, Kurokawa, K., and Nakayama, H. (2014). Chapter 4: Bacterial Lipoproteins: Biogenesis, Virulence/Pathogenicity and Trafficking. p.133-177. In: Bacterial Membranes: Structural and Molecular Biology. Remault, H., Fronzes, R. (eds). Caister Academic Press, Norfolk, UK

2013 **Brülle, J.K.**, Tschumi, A., and Sander P., (2013). Lipoproteins of Slow-growing Mycobacteria Carry Three Fatty Acids and Are *N*-acylated by Apolipoprotein *N*-acyltransferase BCG_2070c. *BMC Microbiol.* 13:223.

2011 Widdick, D.A., Hicks, M.G., Thompson, B.J., Tschumi, A., Chandra, G., Sutcliffe, I.C., **Brülle, J.K.**, Sander, P., Palmer, T., and Hutchings, M.I. (2011). Dissecting the Complete Lipoprotein Biogenesis Pathway in *Streptomyces scabies*. *Mol. Microbiol*. 80: 1395-1412.

2010 **Brülle, J.K.***, Grau, T.*, Tschumi, A., Auchli, Y., Burri, R., Polsfuss, S., Keller, P.M., Hunziker, P., and Sander, P. (2010). Cloning, Expression and Characterization of *Mycobacterium tuberculosis* Lipoprotein LprF. *Biochem. Biophys. Res. Commun.* 391: 679-684.

** Juliane K. Brülle and Thomas Grau contributed equally to this work*

POSTER PRESENTATIONS

2014 Keystone Symposium on Molecular and Cellular Biology. "Novel Therapeutic Approaches to Tuberculosis", Keystone, Colorado, USA

Analysis of Lipoprotein Modification in Mycobacterium bovis *BCG and identification of BCG_2070c as Functional Apolipoprotein* N-*Acyltransferase*

J.K. Brülle*, A. Tschumi and P. Sander

2011 4th Congress of the Federation of European Microbiological Societies (FEMS), Geneva, Switzerland

Structural Analysis of the Lipoprotein Membrane Anchor from Slow-growing Mycobacteria

J.K. Brülle*, A.Tschumi, Y. Auchli, P. Hunziker and P. Sander

Lipoprotein Localization Signals in Mycobacteria

T. Grau*, A. Tschumi, J.K. Brülle and Peter Sander

Keystone Symposium on Molecular and Cellular Biology. "Tuberculosis Immunology: Cell Biology and Novel Vaccination Strategies", Vancouver, Canada

The Membrane Anchor of Mycobacterial Lipoproteins Has Three Arms

A. Tschumi*, J.K. Brülle, T. Grau, Y. Auchli, P. Hunziker and P. Sander

2010 69th Annual Assembly of the Swiss Society of Microbiology (SSM), Zurich, Switzerland

The Lipid Anchor of Mycobacterial Lipoproteins

J.K. Brülle*, A. Tschumi, T. Grau, Y. Auchli, P. Hunziker and P. Sander

* Presenting author

2009 Annual Meeting of the Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) and 62nd Annual Meeting of the Deutsche Gesellschaft für Hygiene und Mikrobiologie (DGHM), Hannover, Germany

Molecular Characterization of Mycobacterium tuberculosis *Lipoprotein LprF*

J.K. Brülle*, T. Grau, A. Tschumi, Y. Auchli, R. Burri, S. Polsfuss, P.M. Keller, P. Hunziker and P. Sander

Dissecting of the Lipoprotein Synthesis Pathway in Mycobacteria

T. Grau, A. Tschumi, S. Polsfuss, J.K. Brülle, P.M. Keller, P. Selchow, T. Rosenberger, A. Petrera, B. Amstutz and P. Sander*

* Presenting author

CURRICULUM VITAE

PERSONAL DATA

Name:	Brülle
First Names:	Juliane Kristina
Date of Birth:	January 31st, 1983
Place of Birth:	Geseke, Germany
Nationality:	German

EDUCATION & RESEARCH ACTIVITIES

2009-Present	**Ph.D. thesis** at the Institute of Medical Microbiology, University of Zurich, Switzerland, under supervision of Prof. Dr. Peter Sander and Prof. Dr. Erik C. Böttger *Structural Analysis of Lipoprotein Modifications in Mycobacteria and other GC-rich Gram-positive Bacteria*
2008-2009	**Diploma work/ Master thesis** at the Institute of Medical Microbiology, University of Zurich, Switzerland, under supervision of Prof. Dr. Peter Sander *Cloning, Expression, Localization and Characterization of* Mycobacterium tuberculosis *Lipoproteins LprF and LprJ*
2005-2009	**Graduate studies** in biology (Diplom/M.Sc.) at the Albert-Ludwigs-University of Freiburg im Breisgau, Germany
2003-2005	**Undergraduate studies** in biology (Vordiplom) at the Philipps-University of Marburg, Germany
1993-2002	**Secondary school** education (Allgemeine Hochschulreife) at Städtisches Gymnasium Erwitte, Germany